

Dødelighetsmodellering med varierende datagrunnlag

av

LINE BUSET

MASTEROPPGAVE

for graden

Master i modellering og dataanalyse

(Master of science)



*Det matematisk-naturvitenskapelige fakultet
Universitetet i Oslo*

Mai 2014

*Faculty of Mathematics and Natural Sciences
University of Oslo*

Takk

Til veileder Erik Bølviken: For en interessant oppgave, og for Fortran-kode til parameterestimering og bootstrapping. For stimulerende faglige diskusjoner og veiledning når jeg trengte det som mest. Og for inspirerende forelesninger og pensum i STK4540 og STK4500.

Sammendrag

Lee-Carter type modeller, og to-trinns estimering, er standardmetoden i demografi og forsikring, men den kan virke ustabil og følsom når parametrene skal bestemmes fra historiske data. En enkelt pensjonskasse eller privatmarkedet for livsvarige annuiteter er eksempler der et relevant datagrunnlag er begrenset. Lee-Carter modellen er motivert ut fra en meget stor populasjon. Et nærliggende spørsmål er dermed om mindre fleksible parametriske modeller vil være mer stabile, og derfor fremskrive dødeligheten bedre enn Lee-Carter type modeller når den tilgjengelige datamengden er mer beskjeden. Ett-trinns estimeringsmetoder er mer implementeringskrevende, men har vist seg å resultere i mer nøyaktige parameterestimer. Gjennom simuleringseksperimenter og en ex-post analyse av tilpasnings- og fremskrivningsfeil, utforsker denne oppgaven sammenhengen mellom modellfleksibilitet og evne til å håndtere tilfeldig variasjon, og også sammenhengen mellom effekt av ett-trinns estimering og datamengde. En mindre fleksibel parametriske modell introduseres og videreutvikles, og det indikeres at den er overlegen Lee-Carter type modeller i håndtering av tilfeldig variasjon. En ett-trinns estimeringsmetode presenteres, og det antydes at fordelene mer enn oppveier implementeringskravet i situasjoner der datagrunnlaget er begrenset.

Innhold

| | | |
|----------|---|-----------|
| 1 | Innledning | 1 |
| 2 | Metode | 4 |
| 2.1 | Historikk og kriterier for modellseleksjon | 4 |
| 2.2 | Lee-Carters modell | 5 |
| 2.3 | Et parametrisk alternativ | 9 |
| 2.4 | Ett-trinns estimeringsmetode | 9 |
| 2.5 | Tilpasning | 11 |
| 2.6 | Fremskrivning og Monte Carlo | 12 |
| 2.7 | Usikkerhet | 15 |
| 3 | Monte Carlo undersøkelser | 16 |
| 3.1 | Post- mot pre-glatting i Lee-Carters modell | 16 |
| 3.1.1 | Usikkerhet i dødelighetstabellene | 17 |
| 3.1.2 | Økonomisk usikkerhet | 20 |
| 3.2 | Glatting og parametrisering | 22 |
| 3.2.1 | Glatting av Lee-Carters modell | 22 |
| 3.2.2 | Pre-glattet Lee-Carter mot parametrisk | 26 |
| 3.2.3 | Pre-glattet Lee-Carter mot parametrisk: Økonomisk | 31 |
| 4 | Historisk analyse av noen land | 33 |
| 4.1 | Innledning | 33 |
| 4.2 | Et første inntrykk | 38 |
| 4.3 | Noen utvalgte eksempler | 41 |
| 4.3.1 | Sveitsiske menn, scenario 1 | 42 |
| 4.3.2 | Franske menn, scenario 2 | 42 |
| 4.3.3 | Islandske kvinner og menn, scenario 2 | 44 |
| 4.4 | Oppsummering | 48 |
| 5 | Konklusjon og videre arbeid | 49 |
| | Appendiks | 55 |
| A | Øvre grense dødelighetsnedgang | 55 |
| B | Estimeringsmetode: Tekniske detaljer | 55 |
| B.1 | Deterministisk filtrering | 55 |
| B.2 | Kalman-filteret | 57 |
| B.3 | Numerisk optimering | 59 |
| B.4 | Startverdier og post-glatting | 60 |
| C | HMD og initiell databehandling | 60 |

| | | |
|----------|--|-----------|
| D | R kode | 61 |
| D.1 | Initiell databehandling | 61 |
| D.2 | Avsnitt 3 | 62 |
| D.2.1 | Post-glatting av a_x | 62 |
| D.2.2 | Fremskrivninger med konfidensintervall | 63 |
| D.2.3 | Dynamisk engangspremie | 66 |
| D.3 | Avsnitt 4 | 68 |

1 Innledning

Vi lever stadig lenger, men hvor lenge avhenger blant annet av hvor vi bor, kjønn og sosioøkonomisk status. Forventet levealder er høyere for kvinner enn menn, og også for den som er forsikret. I en by som Oslo er det store geografiske forskjeller, en mann bosatt i Vestre Aker bydel er forventet å leve rundt ni år lenger enn en mann på Sagene (Berntsen (2013)). At vi stadig lever lenger, viser også at dødeligheten er en dynamisk prosess, data fra mange år tilbake representerer kanskje ikke nåtidens trend i dødelighetsnedgang. Disse eksemplene illustrerer et typisk dilemma. En stor populasjon (for eksempel et land) gir estimater med lav tilfeldig feil, men er kanskje ikke representativ for den populasjonen vi er ute etter, slik at estimatene blir systematisk skjeve.

Standardmodellen i demografi og forsikring er fortsatt den presentert i Lee and Carter (1992); en ikke-parametrisk dødelighetsmodell som muliggjør fleksibel modellering av aldersspesifikk dødelighetsutvikling. Estimeringsmetoden har imidlertid blitt videreutviklet. Fra tradisjonell to-trinns estimering, parameterestimering etterfulgt av tidsrekke-modellering, med utvikling av første trinn fra singulærverdidekomposisjon (SVD) til likelihoodestimering, se for eksempel Wilmoth (1993) og Brouhns et al. (2002b). Til mer beregningskrevende, men også mer nøyaktig, ett-trinns estimering, se blant annet Czado et al. (2005) eller Pedroza (2006) med Markov Chain Monte Carlo og Jong and Tickle (2006) med kalman-filtrering. Det samme gjelder modellen. Fra en streben mot best mulig tilpasning, og dermed utvikling av mer fleksible versjoner eller varianter av den opprinnelige modellen, se blant annet Booth et al. (2002) eller Renshaw and Haberman (2003a) med andre ordens ledd og Renshaw and Haberman (2006) med kohort ledd. Til mer vekt på fremskrivning, og dermed mindre fleksible glattede versjoner, som i for eksempel Delwarde et al. (2007a) eller Currie (2013). Selv om det ser ut til at vekten har skiftet fra fleksibilitet til stabilitet, er det kun ytterst få bidrag som tematiserer Lee-Carter modellens krav til størrelse på datagrunnlaget, se Zhao (2012) for et eksempel på et unntak.

Finansinstitusjonene i Norge brukte inntil for noen år siden den statiske dødelighets-tariffen K1963. Selv om K1963 er basert på faktisk dødelighet i tidsrommet 1955-1963, og forventet levealder i Norge har økt med rundt 5 år siden det, er tariffens sikkerhetstillegg så betydelige at marginene først nylig er spist opp, se Bølviken and Moe (2008). Kombinasjonen av en for svak tariff og finanskrisen, med etterfølgende rentefall, har kostet livsselskapene dyrt. Som et tiltak fastsatte Finanstilsynet i 2013 et minimumskrav til dødelighetsgrunnlag i kollektiv pensjon for alle norske livsforsikringsselskaper og pensjonskasser, K2013. Startdødeligheten i K2013 er basert på dødelighetserfaringene fra 2005-2009 blant kollektivt pensjonsforsikrede i Finans Norges medlemselskaper, og fremskrivninger baseres på SSBs mellomalternativ, se Baltzersen and Steffensen (2013). SSBs mellomalternativ er basert på en fleksibel versjon av Lee-Carter modellen, inkluderer andre ordens ledd for å optimalisere tilpasningen, og tradisjonell to-trinns estimering, se Keilman and Pham (2005). For å ta hensyn til sosioøkonomiske forskjeller mellom de enkelte pensjonskassene, og også at dødelighetsnedgangen i den forsikrede bestanden forventes å bli høyere enn i den generelle befolkningen, legger Finanstilsynet, uten videre forklaring, til en 12% og 10% margin på henholdsvis startdødeligheten og SSBs mellomalternativ, se Baltzersen and Steffensen (2013). Innføringen av K2013 som minimumtariff er omstridt, og i aktuar-

miljøet diskuteres det om det ikke bør være mulig å skreddersy beregningsgrunnlaget til den enkelte pensjonskassen. En diskusjon som nok en gang grunner i estimeringsproblemet introdusert innledningsvis.

Lavt rentenivå, urolige aksjemarkeder, modernisert folketrygd og antakeligvis også oppreserveringskrav tilknyttet innføring av K2013 er faktorer som bidrar til at bedrifter skifter fra ytelses- til innskuddsbasert pensjon. Ifølge Finans Norges hjemmesider har til sammen 4343 ytelsesbaserte ordninger blitt omdannet til innskuddspensjon i perioden 2002-2012, og antall omdannelser pr år har vært stigende. Overgangen fra ytelse til innskudd og den stadige økningen i forventet levealder er to av hovedargumentene i Lin and Cox (2005) for at det kan forventes en kraftig vekst i det private markedet av livsvarige annuiteter. Økt etterspørsel etter livsvarige annuiteter vil igjen føre til et økt behov for lavkostnadsverktøy for håndtering av risikoen assosiert med langt liv. Reforsikring er det tradisjonelle verktøyet, men Lin and Cox (2005) spår også sterk vekst i det relativt unge markedet av dødelighetsobligasjoner eller mortality linked securities (MLS), se for eksempel Lin and Cox (2005) eller Dowd et al. (2006) for en innføring i MLS og en gjennomgang av det eksisterende markedet.

Vekst i markedet av både private, livsvarige annuiteter og MLS forutsetter rettferdig prising. Prising basert på tariffer med lite forankrede sikkerhetsmarginer, som K2014, vil nok ikke være tilfredsstillende. Prisen på et risikooverføringsobjekt kan generelt uttrykkes som forventet tap pluss en kompensasjon for risikoen, se for eksempel Lane (2000), og for en livsvarig annuitet i alder x og år k som

$$V_{xk} = E(\Pi_{xk})(1 + \delta),$$

der Π_{xk} er engangspremien for alder x i år k og δ representerer risikokompensasjonen. Så hva er V_{xk} ? Dødelighetskomponenten av V_{xk} vil i hvert fall til en viss grad avhenge av samspillet mellom dødelighetsmodell, estimeringsmetode og tilgjengelig datagrunnlag, men i hvilken grad?

En enkelt pensjonskasse og markedet for både private annuiteter og MLS er eksempler på situasjoner der det relevante datagrunnlaget er begrenset, og reiser behovet for best mulige statistiske metoder. Målet med oppgaven er å utforske sammenhengen mellom modellens parametriseringsgrad og modellens evne til å håndtere tilfeldig variasjon, og også sammenhengen mellom effekt av ett-trinns estimering og datamengde. Med parametriseringsgrad menes det her antall parametre og dimensjonen på parameterrommene (om parametrene er begrenset eller ikke). En glattet versjon av Lee-Carter modellen vil i denne sammenheng ha en høyere parametriseringsgrad (et mindre antall parametre) enn en uglattet versjon av Lee-Carter modellen, men en lavere parametriseringsgrad enn en parametrisk modell (alle parametre har begrenset parameterrom) med samme antall parametre. Ett-trinns estimeringsmetoden vi vil se nærmere på i denne oppgaven er beskrevet i avsnitt 2.4. Aktuelle stokastiske dødelighetsmodeller behandles i dette tilfelle som ikke-lineære tilstandsrommodeller, og filtreringsproblemets integraler angripes med kjedet numerisk integrasjon.

Sammenliknet med tradisjonell to-trinns estimering har det vist seg at ett-trinns esti-

mering resulterer i et mer nøyaktig estimat av den latente tidsrekke modellen, og Cairns et al. (2011b) og Bølviken (2013) påpeker at virkningen er størst for små populasjoner der ett-trinns metoden demper betydningen av tilfeldige variasjoner. Hvordan forplanter dette seg til fremskrevet dødelighet og forventet levealder, og hvordan påvirkes V_{xk} ? Oppveier fordelene ved ett-trinns estimering det tilhørende implementeringskravet? For å belyse disse spørsmålene utføres det i avsnitt 3.1 et simuleringseksperiment der virkeligheten er antatt å følge Lee-Carter modellen. Sammenhengen mellom effekt av ett-trinns estimering og størrelse på datagrunnlag utforskes ved å sammenlikne stabiliteten til en glattet versjon av Lee-Carter modellen ved post- og pre-glatting (to- og ett-trinns estimering) med varierende størrelse på datagrunnlaget.

En generell regel er at jo mindre data det er tilgjengelig, dess større er behovet for å begrense modellens matematiske struktur (øke modellens parametriseringsgrad). Glatting av Lee-Carter modellen fører som regel til en mer stabil og regelmessig fremskrivning av dødeligheten, se avsnitt 2.2 med referanser. Det forventes at glattingseffekten øker med minkende datagrunnlag, men i hvilken grad? Et naturlig neste spørsmål er da om Lee-Carter type modeller bør erstattes med mindre fleksible parametriske modeller når datagrunnlaget er begrenset? Gjennom å undersøke disse spørsmålene utforsker avsnitt 3.2 sammenhengen mellom modellens parametriseringsgrad og dens evne til å takle tilfeldig variasjon. I avsnitt 3.2.1 utføres det et simuleringseksperiment der virkeligheten, som i avsnitt 3.1, er antatt å følge Lee-Carter modellen. Størrelsen på datagrunnlaget varierer, og stabiliteten til en uglattet versjon av Lee-Carter modellen sammenliknes med en glattet versjon. I avsnitt 3.2.2 er virkeligheten antatt å følge enten Lee-Carter modellen eller en ny parametriske modell foreslått av Bølviken og videreutviklet her, se avsnitt 2.3 for presentasjon og avsnitt 3.2.2 for videreutvikling. Som i avsnitt 3.2.1 varierer størrelsen på datagrunnlaget, og stabiliteten til det parametriske alternativet sammenliknes med en glattet versjon av Lee-Carter modellen.

I avsnitt 3 studeres stabilitet, hvordan metoden (modell og estimeringsmetode) håndterer tilfeldige variasjoner, under forutsetningen om at valget av modell er korrekt (virkeligheten følger modellen). Men virkeligheten følger ingen kjent modell, og i avsnitt 4 undersøkes virkelige data. Avvik mellom modell og virkelighet kan skyldes både estimeringsfeil og valg av modell (at modellen ikke passer). En god tilpasning er heller ikke ensbetydende med en god fremskrivning, og i praksis er det metodens fremskrivningsevne som betyr noe (gir økonomiske konsekvenser). Hvis datagrunnlaget er begrenset, kan en modell fange opp trender i datasettet som ikke er representative for den underliggende populasjonen, men som grunner i tilfeldige variasjoner. Resultatet er kanskje en god tilpasning til datasettet, men ikke til den underliggende populasjonen, og den langsiktige dynamikken under modellen vil ikke nødvendigvis være biologisk rimelig. Vil glatting eller parametrisering (økt parametriseringsgrad) forhindre at modellen fanger opp tilfeldige trender, eller vil det bare føre til en generelt dårligere tilpasning? Og hvilken effekt har ett-trinns estimering på modellens evne til å tilpasse seg historien? Gjennom en ex-post (i ettertid) analyse av tilpasnings- og fremskrivningsevne (se avsnitt 2.7 for forklaring), basert på datasett av ulik størrelse, belyser avsnitt 4 ikke bare disse spørsmålene, men også hvordan det optimale balansepunktet mellom modellstabilitet og -fleksibilitet avhenger av størrelsen på datagrunnlaget.

2 Metode

2.1 Historikk og kriterier for modellseleksjon

Dødelighetsmodelleringens historie er lang, og mange modeller har blitt foreslått siden Gompertz presenterte sin dødelighetsmodell i 1825, se for eksempel Pitacco et al. (2009) for en gjennomgang. Dødelighetsfremskrivning har en kortere historie. Inntil nylig ble dødeligheten behandlet som en deterministisk prosess. Fremskrivningsmetodene i bruk var relativt enkle og involverte ofte subjektive vurderinger, se for eksempel Booth et al. (2006) eller Pitacco et al. (2009). Lee and Carter (1992) markerer overgangen til stokastisk dødelighetsmodellering og fremskrivning. En stokastisk dødelighetsmodell behandler observert dødelighet, eller observert antall døde, som utfall av stokastiske variabler. Videre hensyntas usikkerhet i fremtidig dødelighet ved å foreta fremskrivninger ved bruk av stokastiske prosesser. Resultatet av en fremskrivning er dermed ikke bare et punkttestimat, men en fremskrivningsfordeling, og relevante størrelser som kvantiler er på denne måten tilgjengelige.

I Cairns et al. (2006a) og Cairns et al. (2008) foreslås det en rekke kvantitative og kvalitative kriterier som en god stokastisk dødelighetsmodell bør oppfylle. Eksempler på kvantitative kriterier er at modellen bør være konsistent med historiske data og at parameterestimater og modellfremskrivninger bør være robuste med hensyn til tilpasningsperiode og aldersgruppe. Cairns et al. (2009) og Dowd et al. (2010) bruker blant annet disse kriteriene for å foreta en ex-post sammenlikning av forskjellige modeller. Cairns et al. (2009) bruker også kvalitative kriterier som vanskelighetsgrad på implementering, antall parametre, om metoden hensyntar usikkerhet i parameterestimater og om modellen tar høyde for mulige kohort-effekter. Cairns et al. (2011a) foretar en ex-ante sammenlikning av fremskrivningsevne med vekt på biologisk rimelighet, et konsept introdusert i Cairns et al. (2006a). Cairns et al. (2011a) evaluerer troverdigheten til modellfremskrivningene ved å vurdere om den langsiktige dynamikken under modellen er biologisk rimelig, og om fremskrivningsfordelingen er troverdig og konsistent med historien.

Felles for de nevnte studiene, og dominerende trend for sammenlikningsstudier innen dødelighetsmodellering, er at de evaluerer prestasjonen til Lee-Carter-baserte metoder på store populasjoner. En meget relevant, men lite utforsket, problemstilling er hvordan dagens metoder fungerer på mindre porteføljer. Listen over modellkriterier i Cairns et al. (2008) adresserer ikke denne problemstillingen, og den kunne kanskje utvides til å inkludere et punkt om hvor sensitiv metodens prestasjon er til størrelsen på datagrunnlaget. I denne oppgaven er evaluering og sammenlikning av metoder inspirert av utvalgte kriterier fra Cairns et al. (2008), men med et overordnet fokus på hvordan størrelsen på datagrunnlaget påvirker metodenes prestasjon.

I denne sammenheng omformuleres kriteriet om at parameterestimater og modellfremskrivninger bør være robuste med hensyn til tilpasningsperiode og aldersgruppe til at parameterestimater og modellfremskrivninger bør være robuste med hensyn til størrelsen på datagrunnlaget. Oppbygningen av simuleringseksperimentene i avsnitt 3 er inspirert av dette modifiserte kriteriet. Ex-post analysen av tilpasnings- og fremskrivningsevne i avsnitt 4 er motivert av kriteriet om at modellen bør være konsistent med historiske data og

kriteriene om biologisk rimelighet. Hva som er viktigst av robuste parameterestimerer og konsistens med historiske data, avveiningen eller balansepunktet mellom modellstabilitet og -fleksibilitet, avhenger av hva modellen skal brukes til og størrelsen på datagrunnlaget.

2.2 Lee-Carters modell

De amerikanske demografene Ronald D. Lee og Lawrence R. Carter presenterte i 1992 en bilineær modell for å beskrive og fremskrive tidsvarierende aldersspesifikke dødeligheter, se Lee and Carter (1992). Lee og Carter utviklet metoden spesifikt for alle-årsaks dødelighet i den amerikanske populasjonen for perioden 1933-1987. Arbeidet ble godt mottatt, og det amerikanske folketellingsinstituttet har i ettertid brukt modellen som utgangspunkt for sine befolkningsprosjeksjoner, se Hollmann et al. (2000). Metoden er enkel og modellen har en intuitiv fortolkning, og Lee-Carter type modeller har i senere tid blitt brukt til å modellere og fremskrive alle-årsaks så vel som årsaks-spesifikk dødelighet for mange land og tidsperioder, se for eksempel Girosi and King (2008), Pitacco et al. (2009) eller Booth et al. (2006) samt referanser i disse. Tuljapurkar et al. (2000) har for eksempel brukt Lee-Carter modellen på G7 landene. I Norge er Statistisk sentralbyrås dødelighetsfremskrivninger for den norske befolkningen basert på Lee-Carter, se Keilman and Pham (2005). Lee-Carter baserte modeller er i dag internasjonal standard innen både demografi og forsikring, se for eksempel Booth and Tickle (2008).

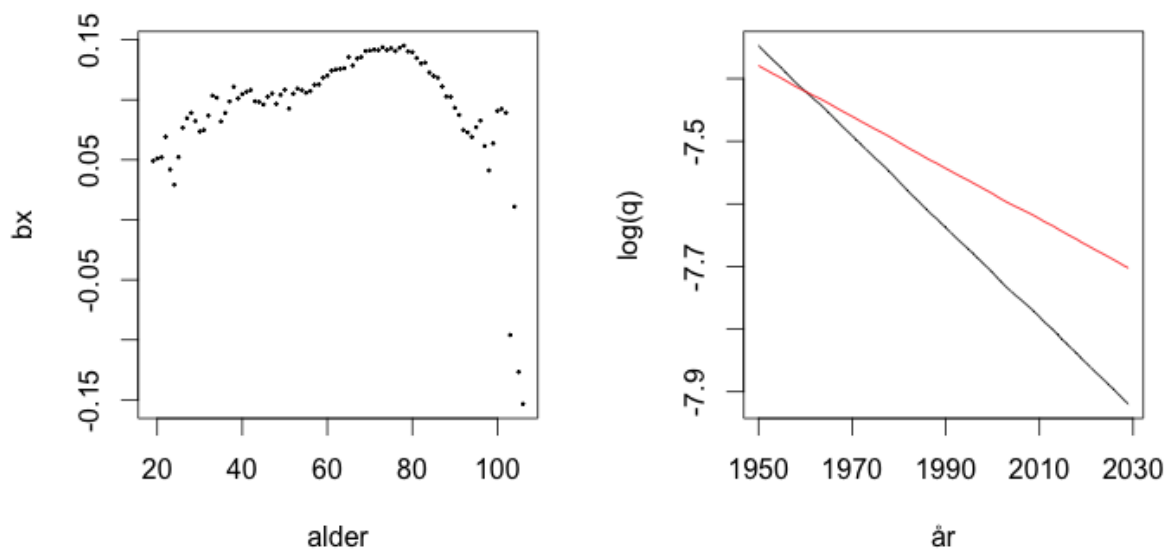
Den opprinnelige Lee-Carter modellen er gitt ved

$$\log(q_{xk}) = a_x + b_x t_k + \epsilon_{xk}.$$

Her er q_{xk} dødeligheten for alder x i år k , a_x er gjennomsnittlig $\log(q_{xk})$ over tilpasningsperioden, t_k representerer det generelle dødelighetsnivået i år k , b_x beskriver hvor sensitiv dødeligheten i alder x er til forandringer i t_k , og ϵ_{xk} er residualer. b_x og t_k estimeres med singularverdidekomposisjon (SVD), og en unik løsning fås ved å sette summen av b_x til 1 og summen av t_k til 0. Deretter reestimeres t_k slik at det for hvert år er en overensstemmelse mellom faktisk og tilpasset total antall døde. For fremskrivningsformål tilpasses en tidsrekkemodell til t_k . Ved bruk av Box-Jenkins metode valgte Lee og Carter en tilfeldig gang med drift

$$t_k = t_{k-1} + \zeta + \sigma e_t,$$

der $e_t \sim N(0,1)$. Det har i ettertid vist seg at t_k i de aller fleste tilfeller følger en tilfeldig gang med drift, se for eksempel Tuljapurkar et al. (2000), Lee and Miller (2001) eller Booth and Tickle (2008). De aldersspesifikke dødelighetene og funksjoner av disse fremskrives ved bruk av estimert a_x og b_x og fremskrevet t_k . Lee and Carter (1992) viser at fremskrivningsfeilen i det lange løp domineres av usikkerhet i fremskrivning av t_k , og velger derfor å se bort fra feil grunnet usikkerhet i parameterestimering. Det har senere vist seg at fremskrivningsfeil grunnet parameterestimering bør hensyntas og at konfidensintervallene presentert i Lee and Carter (1992) dermed er for smale, se for eksempel Pedroza (2006) eller Brouhns et al. (2005).



Figur 1: b_x estimer basert på sveitsiske kvinner 1950-1999 alder 18-105 (til venstre), og fremskrivninger av $\log(q_{23,k})$ (i rødt) og $\log(q_{24,k})$ (i svart) fra første baseår, 1950, og 80 år frem i tid (til høyre).

Lee og Carter tilpasset sin opprinnelige modell i to trinn; parameterestimering ved SVD etterfulgt av tidsrekkemodellering av t_k . Siden har det blitt eksperimentert mye med trinn en av tilpasningsprosessen. Wilmoth (1993) presenterte en vektet minste kvadraters løsning, og utviklet også en likelihoodløsning. Likelihoodbasert estimering etterfulgt av tidsrekkemodellering av t_k har senere blitt tatt i bruk av blant annet Brouhns et al. (2002b), Delwarde et al. (2007a), Delwarde et al. (2007b), Zhao (2012) og Currie (2013). De to trinnene kan kombineres med tilstandromformuleringer i en likelihoodsetting eller ved bruk av bayesiansk metodologi. Begge disse ett-trinns metodene resulterer i et mer nøyaktig estimat av den latente tidsrekke modellen (Cairns et al. (2011b)). Cairns et al. (2011b) og Bølviken (2013) påpeker at denne forskjellen mellom ett- og to-trinns estimering er størst for små populasjoner der ett-trinns metoden demper betydningen av tilfeldig variasjon. Praktiske eksempler inkluderer Czado et al. (2005) og Pedroza (2006) med Markov Chain Monte Carlo (MCMC), Jong and Tickle (2006) med kalman-filtrering og Bølviken (2013) med kjedet numerisk integrasjon (se avsnitt 2.4).

En svakhet ved Lee-Carter modellen er at b_x -estimatene som oftest har et uregelmessig mønster som forplanter seg til estimert og fremskrevet dødelighet, se for eksempel Renshaw and Haberman (2003b), Girosi and King (2008), Delwarde et al. (2007a) eller Currie (2013). Figur 1 viser et eksempel basert på sveitsiske kvinner. Selv med 50 år med data og relativt mange personer pr år har b_x -estimatene et såpass uregelmessig mønster at fremskrivningene av eksempelvis $\log(q_{25,k})$ og $\log(q_{26,k})$ krysser hverandre allerede etter 12 år. At dødeligheten skal divergere på denne måten fra en alder til den neste er urealistisk, og for fremskrivningsformål bør b_x -estimatene derfor glattes.

Delwarde et al. (2007a) viser at glattingen av b_x bør foretas i og ikke etter estimeringsfasen, slik at estimatene av a_x og t_k tilpasser seg de nye b_x -estimatene, og glatter b_x med b-splines ved bruk av penalized-loglikelihood. Currie (2013) bruker penalized-GLM og glatter både a_x og b_x med b-splines. Zhao (2012) glatter a_x og b_x med kubiske splines og andre additive funksjoner. I en bayesiansk setting foretas glatting ved å spesifisere hensiktsmessige priorfordelinger, se for eksempel Czado et al. (2005) eller Pedroza (2006). Jong and Tickle (2006) skreddersyr tilstandsromrammeverket slik at a_x og b_x glattes med b-splines. Bølviken (2013) glatter a_x med Gompertz-Makeham modellen og b_x med kvadratiske splines.

Et annet ofte kritisert aspekt ved Lee-Carter modellen er antakelsen om tidsinvariant b_x . Lee and Miller (2001) viser at aldersmønsteret i dødelighetsnedgangen i USA og andre i-land har forandret seg systematisk i løpet av det 20 århundret. I perioden 1900-1950 var den gjennomsnittlige dødelighetsnedgangen per år høyere for aldersgruppen 0-50, mens det i perioden 1950-2000 observeres omtrent lik gjennomsnittlig dødelighetsnedgang for alle aldre over 15 år. Lee and Miller (2001) foreslår derfor å basere fremskrivninger på en tilpasningsperiode som starter i 1950, som Tuljapurkar et al. (2000), og å anta konstant b_x i denne perioden. Booth et al. (2002) viser at aldersmønsteret i dødelighetsnedgangen for Australia er ustabil også i perioden 1950-2000, og foreslår en metode for å velge den tilpasningsperioden som best oppfyller antakelsen om tidsinvariant b_x .

Lee-Carter modellen har siden publisering i 1992 blitt modifisert og videreutviklet. Bell (1997) og Lee and Miller (2001) viser at modellens fremskrivningsevne forbedres ved å sette a_x til logaritmen av observert dødelighet i siste baseår. Dette er nå bredt akseptert som standard metode, se for eksempel Booth and Tickle (2008). Booth et al. (2002) og Renshaw and Haberman (2003a) utvider den opprinnelige modellen til å inkludere høyere ordens ledd for i større grad å fange opp alder-tid interaksjoner. Renshaw and Haberman (2006) utvider modellen til å inkludere kohort-effekter. Resultatet av begge utvidelsene er sannsynligvis en bedre tilpasning, men også en større risiko for å fange opp og frem-skrive tilfeldige eller flyktige trender, se for eksempel Lee and Miller (2001). Modeller som inkluderer kohort-effekter har i tillegg et begrenset bruksområdet ettersom de krever en meget lang tilpasningsperiode, se for eksempel Pitacco et al. (2009).

Lee-Carter modellen er motivert ut fra et meget stort land. Lee and Carter (1992) observerte at fremskrivningene er ustabile når baseperioden er så kort som 10-20 år. Begrenset datagrunnlag i baseårene forsterker denne ustabiliteten, se Zhao (2012), Richards and Currie (2009) eller Jarner and Kryger (2011) for praktiske eksempler. Zhao (2012) presenterer en mindre fleksibel Lee-Carter type modell, med mer struktur, som takler datasett med få baseår og begrenset datagrunnlag pr år. Zhao (2012) bruker kubiske splines og andre additive funksjoner til å approksimere a_x og b_x og bruker tilnærmingen $k_t = t$. Modellen formuleres som en logistisk regresjon, og estimeringsproblemet reduseres til en utfordring innen modellvalg der funksjoner, enderestriksjoner og optimale skjøter velges ved bruk av residual deviance.

I denne oppgaven vil vi se nærmere på metoden presentert i Bølviken (2013). Bølviken

(2013) gir følgene ikke-lineære tilstandsromformulering av Lee-Carter modellen

$$q_{xk} = \frac{e^{a_x - b_x t_k}}{1 + e^{a_x - b_x t_k}} \quad (1)$$

$$t_k = t_{k-1} + N(\zeta, \sigma^2) \quad \text{og} \quad t_0 = 0. \quad (2)$$

Denne formuleringen av q_{xk} er også brukt i blant annet Zhao (2012) og Cairns et al. (2006b), og sikrer sannsynligheter mellom 0 og 1 for alle parameterverdier. Som Pedroza (2006) antar Bølviken (2013) at t_k følger en tilfeldig gang med drift. Han begrunner valget med at det er ønskelig at modellen skal kunne takle korte baseperioder og brukes til fremskrivninger med lengre horisont. Korte baseperioder krever at den latente variabelen modelleres ved en enkel tidsrekkemodell, og det er heller ikke noen grunn til å regne med at detaljene til en mer sofistikert modell skal holde seg stabile i fremtiden. Residualene, ϵ_{xk} , i Lee and Carter (1992) er også utelatt her ettersom det ofte har vist seg at disse er av marginal betydning. Videre modellerer Bølviken (2013) a_x med Gompertz-Makeham modellen slik at

$$a_x = \log \left(\frac{q_{x0}}{1 - q_{q0}} \right) \quad \text{der} \quad q_{x0} = 1 - \exp(-\theta_0 - \theta_1 e^{\theta_2(x-x_0)}), \quad (3)$$

der θ_0 , θ_1 og θ_2 er positive parametre med begrensninger, og x_0 er den første aldersgruppen som skal inkluderes i tilpasningen. Den andre aldersprofilen b_x er mer varierende, og som Zhao (2012) bruker Bølviken (2013) mer fleksible spline funksjoner til å modellere denne. Bølviken (2013) bruker kvadratiske splines med n_s jevnt fordelte vendepunkter. Hvis n_x er total antall aldersgrupper er $x_s = sn_x/n_s$ endepunktet til segment s for $s = 1, \dots, n_s$. Funksjonen kan gjøres kontinuerlig med kontinuerlig første derivert, men med diskontinuerlig andre derivert i endepunktene x_s . Det er ikke uvanlig å anta $q_{xk} = q_x$ for veldig høye aldersgrupper, med andre ord en konstant maksimal alder, se for eksempel Pitacco et al. (2009). Bølviken (2013) formaliserer dette ved å sette

$$b_x = 0 \quad \text{og} \quad \frac{db_x}{dx} = 0 \quad \text{i} \quad x = x_{n_s}.$$

Den resulterende modellen for b_x kan formuleres som

$$b_x = c_1(x_1 - x)_+^2 + \dots + c_s(x_{n_s} - x)_+^2 \quad (4)$$

der $z_+ = z$ hvis $z > 0$ og 0 ellers og c_1, \dots, c_{n_s} er koeffisienter med en valgt begrensning på maksimum absolutt verdi. Bølviken (2013) normaliserer modellen ved å sette

$$\sum_{x=1}^{n_x} b_x^2 = 1.$$

Resultatet er en litt mindre fleksibel, semiparametrisk, versjon av Lee og Carters opprinnelige modell der antall parametre er redusert fra $2n_x + 2$ i Lee and Carter (1992) til $3 + n_s + 2$. Ett-trinns parameterestimering utføres ved bruk av kjedet numerisk integrasjon, se avsnitt 2.4. Alternativt kan estimeringen foretas på tradisjonelt vis, likelihood estimering etterfulgt av post-glatting, se appendiks B.4.

2.3 Et parametrisk alternativ

En generell regel er at jo mindre data det er tilgjengelig dess større er behovet for å begrense modellens matematiske struktur. Et nærliggende spørsmål er dermed om Lee-Carter type modeller bør byttes ut med noe enklere og mindre fleksibelt når datagrunnlaget er begrenset. Bølviken (2013) foreslår følgende parametriske alternativ

$$q_{xk} = \frac{\xi_x}{\left((1 - \xi_x^{1/\gamma})e^{t_k} + \xi_x^{1/\gamma}\right)^\gamma} \quad (5)$$

$$t_k = t_{k-1} + N(\zeta, \sigma^2), \quad (6)$$

der $t_0 = 0$, γ er et positivt parameter, og $q_{x0} = \xi_x$ modelleres med Gompertz-Makeham slik at

$$\xi_x = 1 - \exp(-\theta_0 - \theta_1 e^{\theta_2 x}), \quad x = 1, \dots, n_x. \quad (7)$$

I motsetning til Lee-Carter modellen sikrer dette parametriske alternativet at q_{xk} er en monoton funksjon av x . Videre vil $q_{xk} \rightarrow \xi_x$ hvis $\gamma \rightarrow 0$ og resultatet er en statisk dødelighetsmodell. Hvis $\gamma \rightarrow \infty$ vil $q_{xk} \rightarrow \xi_x^{e^{t_k}}$ (se appendiks A for utledning), og $\xi_x^{e^{t_k}}$ er dermed en øvre grense for hvor fort dødeligheten synker ettersom t_k vokser. De seks parametrene γ , θ_0 , θ_1 , θ_2 , ζ og σ estimeres med ett-trinns metoden presentert i avsnitt 2.4.

2.4 Ett-trinns estimeringsmetode

Et naturlig rammeverk for stokastisk dødelighetsmodellering er en ikke-lineær tilstandsrommodell hvor tidseffekten er modellert ved hjelp av en underliggende, ikke observerbar, stokastisk prosess, $\{t_k\}$, se Czado et al. (2005), Jong and Tickle (2006) eller Pedroza (2006) for praktiske eksempler. La $\{t_k\}$ være begrenset til første-ordens Markov-prosesser, slik at

$$p(t_0, \dots, t_{n_k}) = p(\mathbf{t}_{0:n_k}) = p(t_0) \prod_{k=1}^{n_k} p(t_k | t_{k-1}),$$

og hvis $\{t_k\}$ modelleres med en tilfeldig gang med drift er

$$p(t_{k+1} | t_k) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(t_{k+1} - t_k - \zeta)/(2\sigma^2)}.$$

Datasettene består av $n_x \times (n_k + 1)$ observasjoner, der n_x er antall aldersgrupper og $(n_k + 1)$ antall år, av y_{xk} døde blant N_{xk} individer, hvor den opplagte fordelingen er binomisk, se for eksempel Zhao (2012) eller Cossette et al. (2007). Alternativer er Poisson som i Brouhns et al. (2002b), Czado et al. (2005) og Cossette et al. (2007), og negativ binomisk som i Delwarde et al. (2007b) og Li et al. (2009). La $q_{xk} = g_x(t_k)$ der $g_x(t)$ er en funksjon med i praksis ukjente parametre. Tilstandsvariabelen, t_k , observeres dermed indirekte gjennom $\{q_{xk}\}$. La videre $\mathbf{y}_k = (y_{0k}, \dots, y_{n_x k})$ og $\mathbf{Y}_{0:k} = (\mathbf{y}_0, \dots, \mathbf{y}_k)$. La

$\mathbf{Y}_{0:k}$ være betinget uavhengige observasjonsvektorer gitt $\mathbf{t}_{0:k}$, og \mathbf{y}_k betinget uavhengige observasjoner gitt t_k . Med $p(y_{xk}|t_k) \sim \text{binomisk}(q_{xk}, N_{xk})$ er $p(\mathbf{y}_k|t_k)$ gitt ved

$$p(\mathbf{y}_k|t_k) = \prod_{x=1}^{n_x} \binom{N_{xk}}{y_{xk}} (q_{xk})^{y_{xk}} (1 - q_{xk})^{N_{xk} - y_{xk}}.$$

Likelihoodfunksjonen, \mathcal{L} , er gitt ved

$$\mathcal{L} = p(\mathbf{Y}_{0:n_k}) = \prod_{k=0}^{n_k} p(\mathbf{y}_k|\mathbf{Y}_{0:k-1}),$$

og på logaritmisk form ved

$$\ell = \log \mathcal{L} = \sum_{k=0}^{n_k} \log(C_k), \quad C_k = p(\mathbf{y}_k|\mathbf{Y}_{0:k-1}).$$

Målet er å finne $\log(C_0), \dots, \log(C_{n_k})$ ved bruk av ikke-lineær filtrering for så å estimere modellparametrene ved å maksimere \mathcal{L} . Optimeringen blir i dette tilfellet utført ved bruk av quasi-Newton-metoden, se appendiks B.3.

Et filter er en rekursiv algoritme for å prosessere observasjoner med støy og produsere estimater på den underliggende, ikke observerbare tilstanden til systemet. Fra et bayesiansk synspunkt evaluerer filteret den betingede sannsynlighetsfordelingen til tilstandsvektoren gitt observasjonene tilgjengelig på dette tidspunktet, og oppdaterer denne når en ny observasjon blir tilgjengelig. Mer formelt kan filtreringsproblemet formuleres som to rekursive steg for å beregne $p(t_k|\mathbf{Y}_{0:k})$, se for eksempel Kitagawa (1987), Cressie and Wikle (2011) eller Bølviken and Storvik (2001). Prediksjonssteget bruker tilstandsmodellen til å foreta en a priori, før observasjonen for nåværende steg er tilgjengelig, prediksjon av den betingede sannsynlighetsfordelingen til t_k basert på filtreringsfordelingen fra tidspunkt $k-1$. I oppdateringssteget blir apriorifordelingen korrigert basert på den nye observasjonen, og resultatet er a posteriori filtreringsfordelingen $p(t_k|\mathbf{Y}_{0:k})$.

Under de gitte forutsetningene om betinget uavhengighet, med $p(t_{k+1}|t_k)$, $p(\mathbf{y}_k|t_k)$ og q_{xk} som i likning (1)-(4) eller (5)-(7) som input, og ved bruk av Bayes-setning og annen grunnleggende sannsynlighetsteori, kan filtreringsproblemet for $\{t_k\}$ formuleres som en to-fase rekursiv integrasjonsalgoritme

Prediksjon

$$p(t_k|\mathbf{Y}_{0:k-1}) = \int_{-\infty}^{\infty} p(t_k|t_{k-1})p(t_{k-1}|\mathbf{Y}_{0:k-1})dt_{k-1} \quad (8)$$

Oppdatering

$$p(t_k|\mathbf{Y}_{0:k}) = p(\mathbf{y}_k|t_k)p(t_k|\mathbf{Y}_{0:k-1})C_k^{-1},$$

der

$$C_k = \int_{-\infty}^{\infty} p(\mathbf{y}_k|t_k)p(t_k|\mathbf{Y}_{0:k-1})dt_k,$$

se for eksempel Kitagawa (1987) eller Cressie and Wikle (2011). I dette tilfellet er det C_k vi er ute etter. Rekursjonen starter med oppdatering ved $k = 0$ med priorfordelingen $p(t_0|\mathbf{Y}_{0:-1}) = p(t_0)$ som input og fortsetter til $k = n_k$, og C_0, \dots, C_{n_k} produseres underveis.

Det er generelt vanskelig og som regel umulig å finne analytiske løsninger til filtreringsproblemets integraler, og gode tilnærminger er derfor nødvendig. Mulighetene er mange, se for eksempel Arulampalam et al. (2002) for en gjennomgang og sammenlikning av de sekvensielle metodene som nå vil bli nevnt.

En hyppig brukt metode er det utvidede kalman-filteret, presentert for annet bruk i appendiks B.2. Det utvidede kalman-filteret lineariserer en ikke-lineær tilstandsrommodell slik at det lineære kalman-filteret kan benyttes. Filtreringsfordelingen blir dermed approksimert med en normalfordeling, og det er umulig å vite hvor god denne tilnærmingen til den eksakte fordelingen er.

En annen mulighet er partikkel filtrering. Partikkel filtrering approksimerer filtreringsproblemets integraler direkte, enten deterministisk ved bruk av kjedet numerisk integrasjon, se for eksempel Kitagawa (1987), Bølviken and Storvik (2001) eller Bølviken (2013), eller stokastisk med sekvensiell Monte Carlo, se for eksempel Andrieu et al. (2001) eller Arulampalam et al. (2002). I motsetning til det utvidede kalman-filteret vil partikkel filtrering, med et tilstrekkelig antall partikler (punkter eller trekninger), tendere mot det eksakte estimatet, se for eksempel Andrieu et al. (2001). Valget mellom deterministisk og stokastisk partikkelfiltrering avhenger av dimensjonen til tilstandsrommodellen. For høy-dimensjonale problemer kan sekvensiell Monte Carlo være foretrukket ettersom kjedet numerisk integrasjon kan bli beregningskrevende og vanskelig å implementere (Doucet et al. (2001)). For lav-dimensjonale tilstandsrommodeller, som modellene vi er interessert i, er derimot deterministisk partikkelfiltrering raskere og mer nøyaktig enn den stokastiske versjonen (Bølviken (2013)), se Bølviken and Storvik (2001) for en sammenlikning med sekvensiell Monte Carlo. I vårt tilfelle blir Gauss-Legendre-filtrering brukt for å approksimere filtreringsproblemet for $\{t_k\}$, se appendiks B.1.

Erik Bølviken har utviklet og skrevet et beregningsprogram i Fortran som er basert på den presenterte teorien. Dette programmet blir brukt til parameterestimering. Appendiks B.1, B.2 og B.3 beskriver i grove trekk bakgrunnen for og oppbygningen av estimeringsprogrammets hovedalgoritmer.

2.5 Tilpasning

For å sjekke modellenes tilpasning til datasettene sammenliknes estimert dødelighet \tilde{q}_{xk} og estimert forventet gjenværende levetid \tilde{e}_x med de tilsvarende observerte størrelsene

$$\dot{q}_{xk} = \frac{y_{xk}}{N_{xk}}$$

og

$$\dot{e}_x = \sum_{i=1}^{n_x-x} i \dot{p}_{xk} + \frac{1}{2} \quad \text{med} \quad {}_{j+1}\dot{p}_{xk} = {}_j \dot{p}_{xk} (1 - \dot{q}_{(x+j)k}), \quad (9)$$

der ${}_j\dot{p}_{xk}$ beregnes rekursivt, og er sannsynligheten for at en person med alder x i år k er i live ved alder $x + j$, se for eksempel Gerber (1997). Den dynamiske versjonen av \dot{e}_x fås ved å bytte ut ${}_j\dot{p}_{xk}$ i uttrykket over med ${}_j\dot{p}_{xk} = {}_j\dot{p}_{xk}(1 - \dot{q}_{(x+j)(k+j)})$. Det er ønskelig å bruke en tilpasningsperiode som representerer nåtidens trender, se avsnitt 2.2. Det er derfor ikke mulig å beregne observert forventet levetid dynamisk, ettersom nødvendige data ikke er tilgjengelige. Sammenlikning av observert og estimert forventet levetid baseres derfor på statiske estimater.

For å beregne \tilde{q}_{xk} og \tilde{e}_x ved ett-trinns estimering trengs a posteriori tetthetsfunksjonene $p(t_k|\mathbf{Y}_{0:n_k})$. Hvis $p(t_k|\mathbf{Y}_{0:k})$ og $p(t_{k+1}|\mathbf{Y}_{0:k})$ lagres underveis i filtreringen beskrevet i avsnitt 2.4, kan $p(t_k|\mathbf{Y}_{0:n_k})$ estimeres rekursivt fra $k = n_k - 1$ til $k = 0$ ved

$$p(t_k|\mathbf{Y}_{0:n_k}) = p(t_k|\mathbf{Y}_{0:k}) \int_{-\infty}^{\infty} \frac{p(t_{k+1}|t_k)p(t_{k+1}|\mathbf{Y}_{0:n_k})}{p(t_{k+1}|\mathbf{Y}_{0:k})} dt_{k+1}, \quad (10)$$

se for eksempel Kitagawa (1987) eller Cressie and Wikle (2011) for utledning og avsnitt B.1 for numeriske detaljer. Estimert dødelighet og estimert forventet gjenværende levetid under de gitte modellene for $q_{xk} = g_x(t_k)$ i likning (1)-(4) og (5)-(7) er gitt ved

$$\tilde{q}_{xk} = \int_{-\infty}^{\infty} g_x(t_k)p(t_k|\mathbf{Y}_{0:n_k})dt_k$$

og

$$\tilde{e}_{xk} = \int_{-\infty}^{\infty} \left(\frac{1}{2} + \sum_{i=1}^{n_x-x} {}_i\tilde{p}_{xk} \right) p(t_k|\mathbf{Y}_{0:n_k})dt_k \quad \text{med} \quad {}_{j+1}\tilde{p}_{xk} = {}_j\tilde{p}_{xk}(1 - g_{x+j}(t_k)).$$

For Lee-Carter modellen vil vi også se på tilpasning ved to-trinns estimering, \tilde{q}_{xk} beregnes da som i likning (1) ved bruk av estimert a_x , b_x og t_k , og \tilde{e}_{xk} som i likning (9) ved bruk av \tilde{q}_{xk} .

2.6 Fremskrivning og Monte Carlo

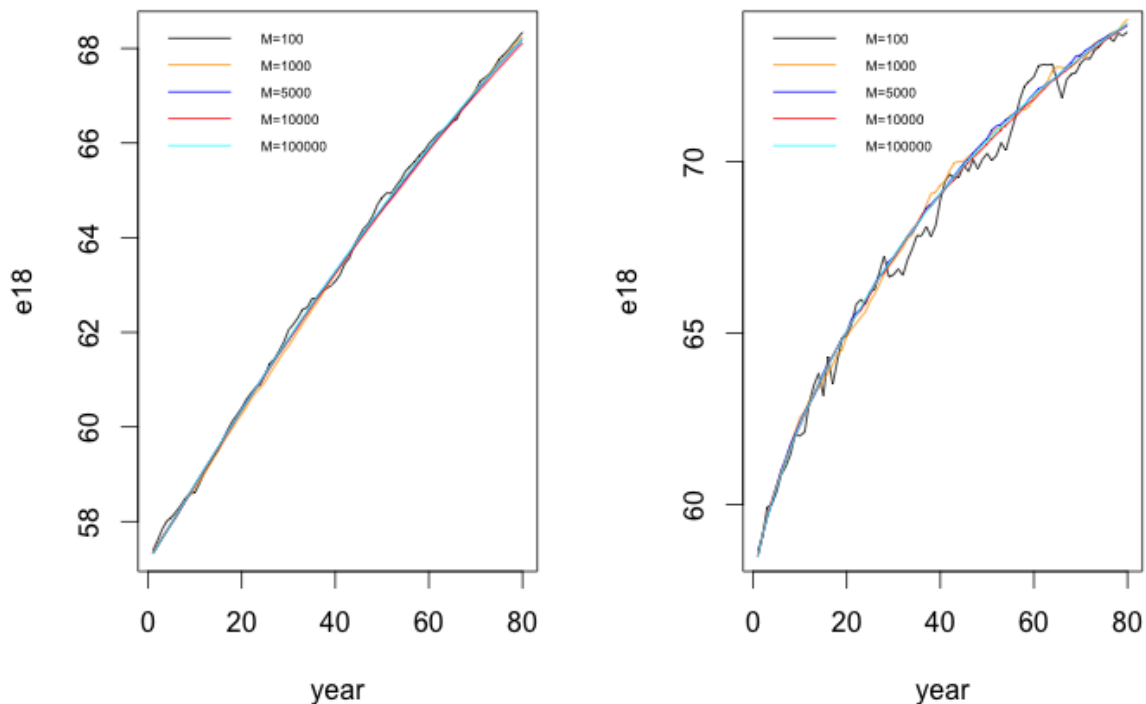
For fremskrivningsformål kan det være hensiktsmessig å foreta estimeringen av modellenes parametre slik at $t_{n_k} = 0$ istedenfor $t_0 = 0$, se avsnitt 2.2. Da er

$$a_x = \log \left(\frac{q_{xn_k}}{1 - q_{xn_k}} \right) \quad \text{og} \quad \xi_x = q_{xn_k}$$

i henholdsvis Lee-Carter og den parametriske modellen. Dette kan, med ett-trinns estimering, oppnås ved å snu datasettene (siste året i tilpasningsperioden settes til år 0 og første til år n_k). Driften, ζ , i tilfeldig gang-modellen for t_k får dermed negativt fortegn, og fremskrivning vil i teorien foregå bakover i tid slik at

$$t_{k-1} = t_k - N(\zeta, \sigma^2).$$

Alternativt kan estimeringen foretas med $t_0 = 0$ og fremskrivningen begynne i $t_{n_k} = \tilde{t}_{n_k}$, se blant annet Booth et al. (2002).



Figur 2: Varierende antall Monte Carlo simuleringer, M . \hat{e}_{18k} (til venstre) og 0.975-kvantilen til $\{e_{18k}^m\}$ (til høyre).

Modellene for q_{xk} er stokastiske. Både forventningsverdien og medianen til fremskrivningsfordelingene, til q_{xk} og funksjoner av disse, er derfor aktuelle punktestimater. Vi velger videre å bruke forventningsverdien. Hyppige brukte variabler for å evaluere fremskrivningsevne er de aldersspesifikke dødelighetene q_{xk} samt forventet levetid e_{xk} , se for eksempel Booth et al. (2006) eller Dowd et al. (2010). Fra et livsforsikringsperspektiv er engangspremien for en livsvarig annuitet i alder x og år k , Π_{xk} , en relevant variabel, se for eksempel Cairns et al. (2011a), Brouhns et al. (2005) eller Brouhns et al. (2002a) og avsnitt 1. e_{xk} og Π_{xk} beregnes fra observert eller fremskrevet dødelighet som i likning 9 og ved

$$\Pi_{xk} = \sum_{i=l}^{n_x-x} d_i^l p_{xk} \quad l = \max(l_r - x, 0), \quad (11)$$

der ${}_{j+1}p_{xk} = {}_j p_{xk}(1 - q_{(x+j)(k+j)})$ eller ${}_{j+1}p_{xk} = {}_j p_{xk}(1 - q_{(x+j)k})$ for henholdsvis dynamisk eller statisk versjon, l_r er pensjonsalderen og $d = 1/r$ hvor r i vårt tilfelle er risikofri rente ettersom vi kun er ute etter dødelighetskomponenten av Π_{xk} eller V_{xk} .

At en modell er overlegen en annen i fremskrivning av de aldersspesifikke dødelighetene behøver ikke nødvendigvis bety at det samme gjelder for aggregerte størrelser, som e_{xk} og

Π_{xk} , og vice versa. Dette grunner i at fremskrivningsfeil for de aldersspesifikke dødelighetene potensielt kan kanselleres mot hverandre i beregningen av de aggregerte størrelsene, og dermed maskere omfanget av fremskrivningsfeil, se Booth et al. (2006) for et praktisk eksempel.

Størrelsene som skal fremskrives er mer eller mindre kompliserte funksjoner av de stokastiske variablene t_k . Monte Carlo metoden er en simuleringsmetode som blant annet kan brukes til å approksimere en sannsynlighetsfordeling dersom analytiske uttrykk er vanskelig eller umulig å finne. Monte Carlo metoden er basert på asymptotisk teori, se for eksempel Cressie and Wikle (2011) eller Ross (2010). Antall simuleringer M som er nødvendig for å oppnå en akseptabel tilnærming, avhenger av hva som anses som akseptabelt, fordelingen som skal tilnærmes og hva den tilnærmede fordelingen skal brukes til. I vårt tilfelle er vi ute etter både forventningsverdi og kvantiler, og det er ikke ønskelig med unødvendig mange simuleringer, ettersom bootstrapbaserte konfidensintervall, se avsnitt 2.7, gjør fremskrivningene beregningskrevende. Monte Carlo estimatene for forventning $E()$ og ϵ -kvantilen $k_\epsilon()$ for en funksjon $g()$ av den stokastiske variabelen eller vektoren X er gitt ved

$$\hat{E}(g(X)) = \frac{1}{M} \sum_{i=1}^M g(X_i)$$

og

$$\hat{k}_\epsilon(g(X)) = g(X_{\epsilon M}),$$

der X_i , $i = 1, \dots, M$, er Monte Carlo realiseringer av X , og $g(X_{\epsilon M})$ er det ϵM elementet i $\{g(X_1) \leq g(X_2) \leq \dots \leq g(X_M)\}$, se Cressie and Wikle (2011) eller Ross (2010). Som eksempel kan punktestimatet til engangspremien for en x -åring i år $z = n_k + 1$, $\hat{\Pi}_{xz}$, estimeres dynamisk ved at det først genereres, for hvert tidssteg $j = 0, \dots, (n_x - x)$, M realiseringer fra tilfeldig gang modellen

$$t_{z+j}^m = t_{z+j-1}^m + N(\zeta, \sigma^2) \quad \text{med} \quad t_{z-1}^m = t_{n_k} = 0 \quad \text{og} \quad m = 1, \dots, M.$$

Deretter estimeres

$$q_{(x+j)(z+j)}^m = g_{(x+j)}(t_{z+j}^m)$$

for $m = 1, \dots, M$ og $j = 0, \dots, (n_x - x)$. Π_{xz}^m beregnes som i likning (11) for $m = 1, \dots, M$ og

$$\hat{\Pi}_{xz} = \frac{1}{M} \sum_{i=1}^M \Pi_{xz}^m.$$

Det eksperimenteres med antall nødvendige simuleringer M . Ettersom tilpasning av de aktuelle modellene som regel gir $\sigma < 0.5$ for aktuelle datasett, anvendes $\sigma = 0.5$. Basert på grafisk evaluering av stabiliteten til de aktuelle estimatene og modellene, for M i området $100 - 100000$, vil det videre bli brukt $M = 1000$ for beregning av punktestimater og $M = 10000$ for beregning av fremskrivningsintervall. Figur 2 viser hvordan punktestimatet og estimert 0.975-kvantil for e_{18k} under Lee-Carter modellen stabiliserer seg ettersom M øker. Tilsvarende resultater oppnås også for den parametriske modellen. Fremskrivningene i avsnitt 3 og 4 foretas i R, se appendiks D for kode.

2.7 Usikkerhet

Usikkerhet i modellfremskrivninger kan kvantifiseres ved bruk av konfidensintervall. Analytiske uttrykk er ikke mulig å finne ettersom to veldig forskjellige kilder til usikkerhet må kombineres: binomisk samplingfeil (tilfeldige feil) i modellparametrene og fremskrivningsfeil i det predikerte tilfeldig gang parameteret t_k . Historisk er det kun fremskrivningsfeil som har blitt hensyntatt, men det har senere vist seg at dette kan resultere i urealistisk smale konfidensintervall, se avsnitt 2.2.

Brouhns et al. (2002a) og Brouhns et al. (2005) presenterer og sammenlikner bruk av parametrisk og semi-parametrisk bootstrapping på en Poisson versjon av Lee-Carter modellen. Koissi et al. (2006) implementerer og anvender residual bootstrapping på den samme modellen. Renshaw and Haberman (2008b) og (2008a) sammenlikner bruk av parametrisk, semi-parametrisk og residual bootstrapping på Poisson versjonen og også på en negativ binomisk og en binomisk versjon av Lee-Carter modellen. Alle disse bootstrap-metodene hensyntar både samplingfeil og fremskrivningsfeil. Alternativet til bootstrapping er å ta i bruk bayesiansk metodologi som i Czado et al. (2005) eller Pedroza (2006).

Simuleringseksperimentene i avsnitt 3 er basert på semi-parametrisk bootstrapping. Semi-parametrisk bootstrapping utføres ved å først generere N bootstrap-realiseringer over antall døde y_{xk}^n , for $x = 1, \dots, n_x$ og $k = 0, \dots, n_k$, der $y_{xk}^n \sim \text{binomisk}(q_{xk}^n, N_{xk})$. Her er q_{xk}^n en realisering fra den opprinnelige tilpassede modellen θ^0 . Deretter reestimeres modellens parametre for hver av de N bootstrap-realiseringene, og resultatet er N bootstrap-modeller $\{\theta^n\}$. Hvis det for eksempel er ønskelig å estimere konfidensintervall rundt punkttestimatet til fremskrevet dødelighet, \hat{q}_{xk} , fremskrives og beregnes punkttestimatet \hat{q}_{xk}^n , som i avsnitt 2.6, for hver av de N bootstrap-modellene θ^n . Øvre grense av konfidensintervallet rundt \hat{q}_{xk} estimeres som ϵ -kvantilen til $\{\hat{q}_{xk}^n\}$, det vil si som det ϵN elementet i $\{\hat{q}_{xk}^1 \leq \hat{q}_{xk}^2 \leq \dots \leq \hat{q}_{xk}^N\}$. Bootstrappingen i avsnitt 3 foretas ved bruk av et Fortran program skrevet av Bølviken, og etterfølgende fremskrivninger utføres i R, se appendiks D for R kode.

Bootstrapping fasiliteter kvantifisering av usikkerhet under en gitt modell, men hensyntar ikke feil grunnet valg av modell. Alle modeller er feil. Modellfeil er generelt vanskelig å kvantifisere, men aktuelle modeller kan sammenliknes. Modellene kan for eksempel tilpasses til deler av et datasett for så å se på overenstemmelse mellom modelltilpasning og observert dødelighet i tilpasningsperioden, og også mellom fremskrevet dødelighet og observert dødelighet i de årene som er utelatt fra tilpasningsperioden. Praktiske eksempler på denne type ex-post analyser inkluderer Lee and Miller (2001), Booth et al. (2006), og Dowd et al. (2010), og også avsnitt 4.

I tillegg til ikke-kontrollerbar usikkerhet grunnet samplingfeil, fremskrivningsfeil, og modellfeil er det også usikkerhet knyttet til bruk av Monte Carlo tilnærminger, se avsnitt 2.6, men ettersom denne usikkerheten kan kontrolleres ved å øke størrelsen på M vil den ikke vies mer oppmerksomhet her.

3 Monte Carlo undersøkelser

3.1 Post- mot pre-glatting i Lee-Carters modell

Ett-trinns estimeringsmetoden presentert i avsnitt 2.4 er relativt implementeringskrevende sammenliknet med den tradisjonelle to-trinns metoden (likelihood estimering etterfulgt av glatting). Bølviken (2013) sammenlikner post- og pre-glatting (to- og ett-trinns estimering) og viser at pre-glatting demper effekten av tilfeldig variasjon og dermed resulterer i mer nøyaktige estimer av b_x og den latente tidsrekke modellen, t_k , i Lee-Carter modellen (likning (1)-(4)). Er denne effekten av pre-glatting såpass betydningsfull at den oppveier vanskelighetsgraden av implementeringen?

For å utforske dette nærmere er det utført et simuleringseksperiment der virkeligheten er antatt å følge Lee-Carter modellen i likning (1)-(4), med parameterverdier basert på populasjonen av sveitsiske kvinner mellom 18-105 år i tidsperioden 1950-1999, se tabell 1. Semi-parametrisk bootstrapping som i avsnitt 2.7, med $N = 500$, θ^0 som i tabell 1 og antall splinesegmenter $n_s = 5$, er utført med både post- og pre-glatting for fire scenarier. Valget av $n_s = 5$ er basert på eksperimenter som indikerer at resultatene er stabile når $5 \leq n_s \leq 10$. Antall år i tilpasningsperioden er variert mellom 50 (scenario 1 og 2) og 20 (scenario 3 og 4) og antall kvinner pr år mellom 1.5 millioner (scenario 1 og 3) og 150 000 (scenario 2 og 4) mens aldersfordelingen er holdt konstant og er lik den for sveitsiske kvinner i årene 1950-1999.

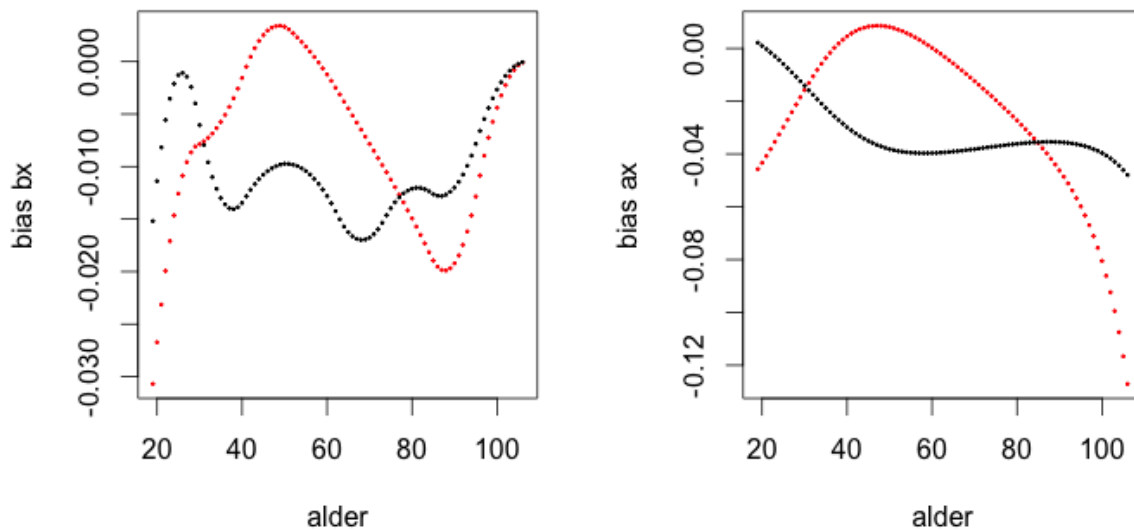
Resultatene i tabell 2 antyder at det er spesielt vanskelig å estimere det dynamiske parameteret σ med stor nøyaktighet. Post-glatting resulterer i en overestimering av σ som er særlig utpreget når antall kvinner pr år reduseres til 150 000, se tabell 2. Pre-glatting fører til mye mindre skjevhet. Av de to aldersprofilene, a_x og b_x , er det b_x som er den vanskeligste å estimere. En versjon av RMSE (root mean square error) for b_x er

$$\text{RMSE}_b = \sqrt{\frac{1}{Nn_x} \sum_{x=1}^{n_x} \sum_{n=1}^N (b_x^n - b_x^0)^2}.$$

Denne størrelsen er tabellert i tabell 3 for de fire scenariene. Pre-glatting fører til mer nøyaktige estimer, særlig for scenario 2 og 3. For scenario 4 er feilen stor uansett. Pre-glatting fører i dette scenariet til underestimering av b_x for rundt 80% av aldersgruppene, mens post-glatting resulterer i underestimering av b_x for alle aldre og total skjevhet er 30 % høyere enn med pre-glatting, se venstre side av figur 3. For scenario 4 fører post-glatting også til underestimering av a_x for alle aldre over 20, se høyre side av figur 3.

Tabell 1: *Lee-Carter modellen tilpasset sveitsiske kvinner.*

| ζ | σ | $10^3\theta_0$ | $10^3\theta_1$ | θ_2 | 10^3c_1 | 10^3c_2 | 10^3c_3 | 10^3c_4 | 10^3c_5 | 10^3c_6 | 10^3c_7 | 10^3c_8 |
|---------|----------|----------------|----------------|------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|
| 0.140 | 0.278 | 0.438 | 0.097 | 0.1108 | 0.1068 | -0.0760 | 0.0175 | -0.1178 | 0.0257 | 0.0646 | -0.0818 | 0.0344 |



Figur 3: *Skjevhet (bias), post- (i svart) og pre-glatting (i rødt) for scenario 4, b_x (til venstre) og a_x (til høyre).*

3.1.1 Usikkerhet i dødelighetstabellene

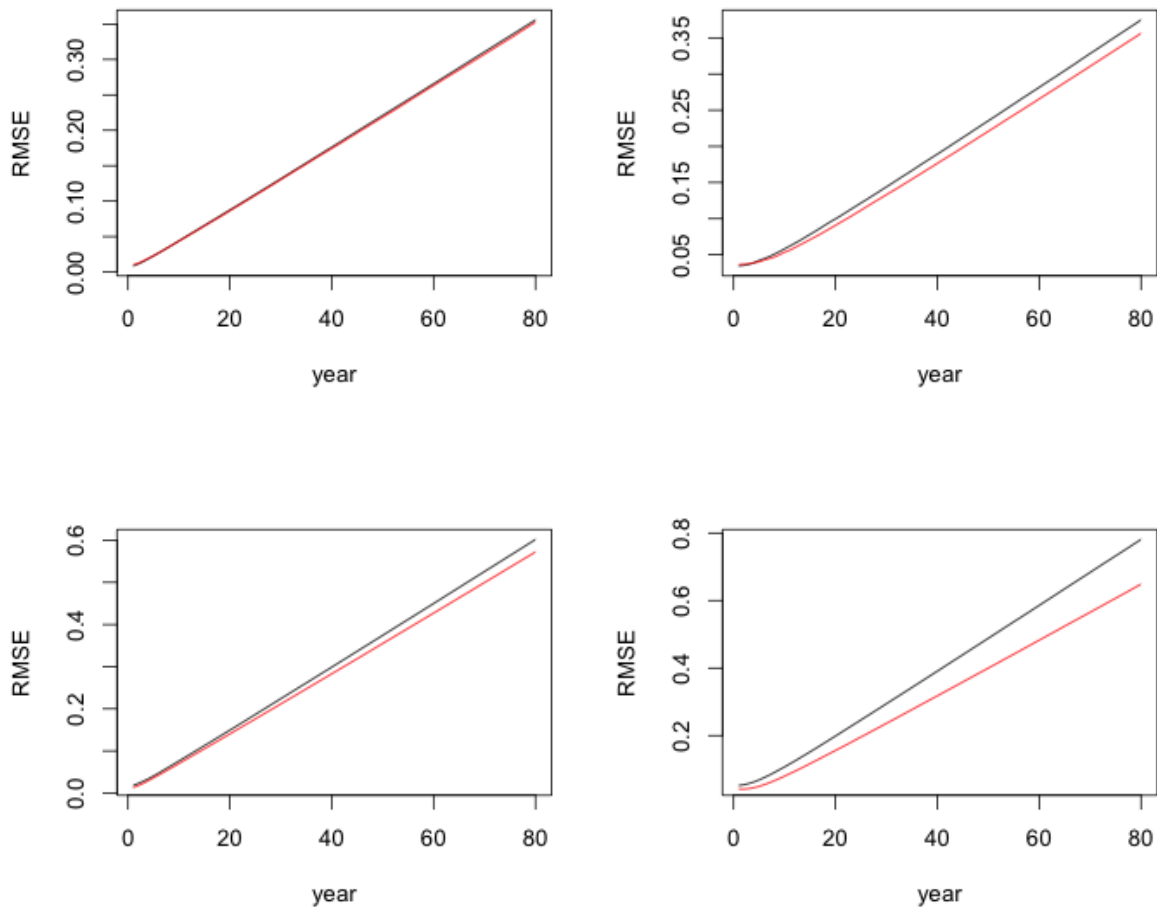
I hvor stor grad forplanter den observerte fordelen med pre-glatting seg til relevante størrelser som $\log(q_{xk})$ og e_{18k} ? Figur 4 viser RMSE pr år for punktestimatet til fremskrevet $\log(q_{xk})$

$$\text{RMSE}_{\log(\hat{q}_k)} = \sqrt{\frac{1}{Nn_x} \sum_{x=1}^{n_x} \sum_{n=1}^N (\log(\hat{q}_{xk}^n) - \log(\hat{q}_{xk}^0))^2},$$

der $\log(\hat{q}_k)$ er summen av $\log(\hat{q}_{xk})$, $x = 1, \dots, n_x$. For scenario 1 er det ikke store forskjeller, men for scenario 2, 3 og spesielt 4 resulterer pre-glatting i mer nøyaktige estimater, og effekten øker med fremskrivningshorisonten.

Tabell 2: *Skjevhet (bias) og standardfeil (se) for tilfeldig gang parametrene for post- og pre-glattet Lee-Carter.*

| | 1.5 mill, 50 år | | | | 150 000, 50 år | | | | 1.5 mill, 20 år | | | | 150 000, 20 år | | | |
|----------|-----------------|--------|---------|--------|----------------|--------|---------|--------|-----------------|--------|---------|--------|----------------|--------|---------|--------|
| | Post | | Pre | | Post | | Pre | | Post | | Pre | | Post | | Pre | |
| | bias | se | bias | se | bias | se | bias | se | bias | se | bias | se | bias | se | bias | se |
| ζ | 0.0018 | 0.0392 | 0.0016 | 0.0388 | 0.0011 | 0.0390 | -0.0017 | 0.0382 | <1e-04 | 0.0651 | -0.0003 | 0.0627 | 0.0077 | 0.0709 | 0.0027 | 0.0636 |
| σ | 0.0128 | 0.0304 | -0.0054 | 0.0316 | 0.1239 | 0.0461 | -0.0062 | 0.0528 | 0.0118 | 0.0483 | -0.0135 | 0.0492 | 0.1778 | 0.1067 | -0.0345 | 0.1103 |

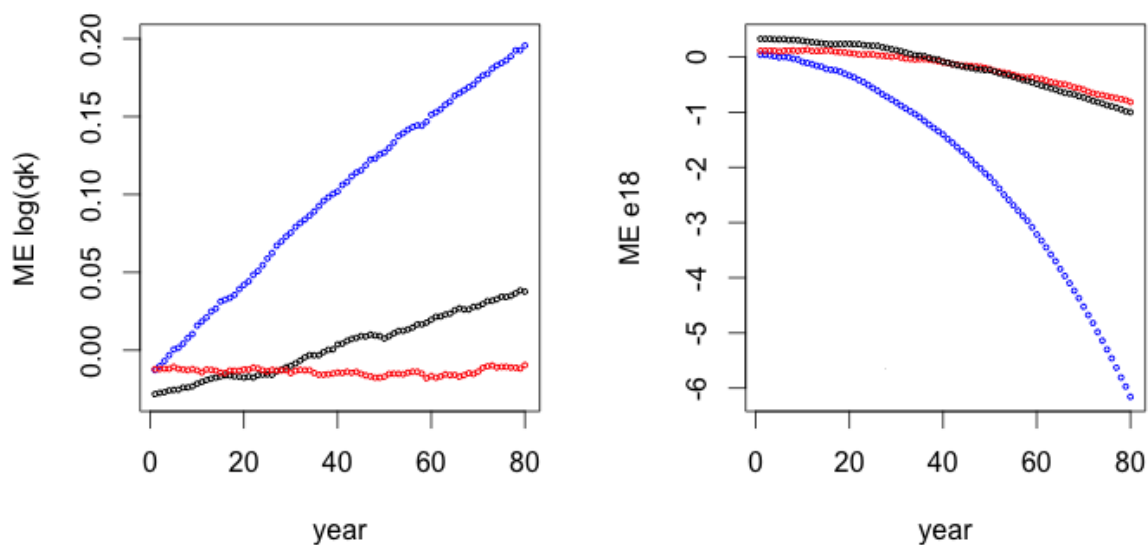


Figur 4: $RMSE_{\log(\hat{q}_k)}$ for post- (i svart) og pre-glattet (i rødt) Lee-Carter, scenario 1 (øverst til venstre), scenario 2 (øverst til høyre), scenario 3 (nederst til venstre) og scenario 4 (nederst til høyre).

For å avdekke eventuell skjevhet for scenario 4 anvendes ME (mean error) pr år

$$ME_{\log(\hat{q}_k)} = \frac{1}{Nn_x} \sum_{x=1}^{n_x} \sum_{n=1}^N (\log(\hat{q}_{xk}^n) - \log(\hat{q}_{xk}^0)).$$

Venstre side av figur 5 viser at post-glatting underestimerer $\log(\hat{q}_k)$ i første del av fremskrivningsperioden. Deretter overestimeres $\log(\hat{q}_k)$ i økende grad med fremskrivningshorisonten. Dette samsvarer med figur 3. For en kort fremskrivningshorisont er $\log(\hat{q}_{xk})$ preget av underestimering av a_x , men etter som fremskrivningshorisonten øker, sørger underestimering av b_x for at skjevhet i $\log(\hat{q}_{xk})$ blir tiltagende positiv. Sammenliknet med post-glatting er skjevhet assosiert med pre-glatting tilnærmet neglisjerbar.



Figur 5: $ME_{\log(\hat{q}_k)}$ for uglattet (i blått), post- (i svart) og pre-glattet (i rødt) Lee-Carter i scenario 4 (til venstre). $ME_{\hat{e}_{18k}}$ for uglattet, post- og pre-glattet Lee-Carter i scenario 4 (til høyre).

Samme trend som for $RMSE_{\log(\hat{q}_k)}$ i figur 4 observeres for RMSE for statisk beregnet forventet gjenværende levetid ved alder 18, $RMSE_{\hat{e}_{18k}}$. Effekten av pre-glatting akselererer med minkende datagrunnlag og etter som fremskrivningshorisonten øker. Høyre side av figur 5 viser at \hat{e}_{18k} overestimeres i første del av fremskrivningsperioden og deretter underestimeres. Bruk av pre-glatting resulterer igjen i mindre skjevhet. Merk at beregningen av \hat{e}_{18k} baseres på \hat{q}_{xk} og ikke $\log(\hat{q}_{xk})$. Det er derfor ingen direkte sammenheng mellom venstre og høyre side i figur 5. For post-glatting kan man likevel se et samsvar i trenden, underestimering av $\log(\hat{q}_k)$ og dermed overestimering av \hat{e}_{18k} i de første årene av fremskrivningsperioden, og motsatt deretter.

Tabell 3: $RMSE_b$ for post- og pre-glattet Lee-Carter.

| 1.5 mill, 50 år | | 150 000, 50 år | | 1.5 mill, 20 år | | 150 000, 20 år | |
|-----------------|--------|----------------|--------|-----------------|--------|----------------|--------|
| Post | Pre | Post | Pre | Post | Pre | Post | Pre |
| 0.0033 | 0.0025 | 0.0199 | 0.0088 | 0.0116 | 0.0083 | 0.0285 | 0.0248 |

3.1.2 Økonomisk usikkerhet

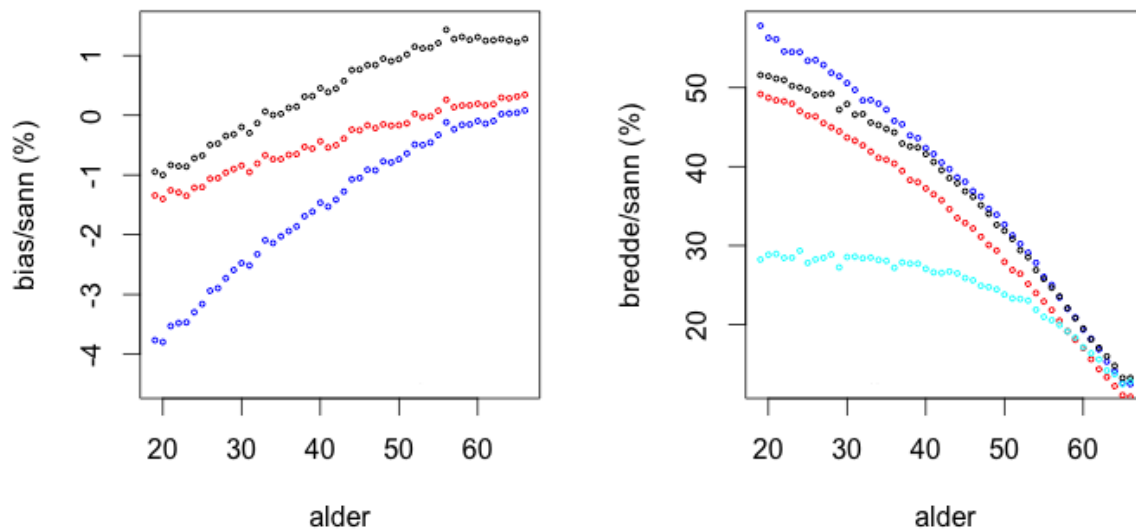
For å kunne si noe om den økonomiske konsekvensen av pre-glatting velger vi å ta for oss dynamisk engangspremie for en 60-åring i scenario 1 og 4. Fra tabell 4 fremgår det at valget mellom post- og pre-glatting ser ut til å være uten betydning med et datagrunnlag på 1.5 millioner pr år og 50 år med data, skjevhet og bredde på konfidensintervall er tilnærmet lik for post- og pre-glatting. Med et datagrunnlag begrenset til 150 000 pr år og 20 år med data er situasjonen en annen, post-glatting fører til over 6 ganger så stor skjevhet og konfidensintervallet er 10% bredere enn ved pre-glatting. Overestimeringen for post-glatting i scenario 4 grunner i underestimering av $\{q_{(x+i)(k+i)}\}$ for $x = 60$, $k = 0$ og $i = 0, \dots, 45$, se trend til venstre i figur 5. En skjevhet på rundt 1.2% høres ikke mye ut, men kan potensielt ha store økonomiske konsekvenser for en bedrift som ønsker å sette bort en livsforsikringsportefølje.

For å utforske dette nærmere beregnes dynamiske engangspremier for alle aldre mellom 18-65. Høyre side av figur 6 viser bredde på konfidensintervall, i prosent av forventningsverdien under modellen i tabell 1, ved bruk av post-, pre- og uglattet Lee-Carter. Her fremkommer det at usikkerhet i parameterestimerer bør hensyntas når datagrunnlaget er såpass begrenset som i scenario 4. Uansett estimeringsmetode er ikke fremskrivningsintervallet i nærheten av å inneholde konfidensintervallet rundt punkttestimatet for aldersgruppen 18-58. Sammenliknet med post-glatting fører pre-glatting imidlertid til rundt 10% mindre usikkerhet i punkttestimatet til Π_{xk} for alle aldre. For et livsforsikringsselskap vil en reduksjon i risiko i denne størrelsesorden antakeligvis oppveie beregningskravet assosiert med pre-glatting.

Fra venstre side av figur 6 fremkommer det at både post- og pre-glatting underestimerer engangspremien for aldersgruppen 18-31 med et gjennomsnitt på henholdsvis 0.6% og 1% av forventningsverdien under modellen i tabell 1. Post-glatting fører i tillegg til overestimering for aldersgruppen 32-65, gjennomsnittlig verdi er på 1.1% mot $< 0.1\%$ ved pre-glatting. En gjennomsnittlig bedrift, med hovedvekten av ansatte over 30 år, som ønsker å sette bort en portefølje av livsvarige annuiteter vil nok, i dette tilfellet, foretrekke bruk av mer beregningskrevende pre-glatting.

Tabell 4: *Dynamisk engangspremien for en 60-åring, $\Pi_{60,0}$. Skjevhet (bias) og bredde på konfidensintervall, verdier for post-glattet Lee-Carter og forholdet mellom post- og pre-glattet Lee-Carter (post/pre).*

| | 1.5 mill, 50 år | | | | 150 000, 20 år | | | |
|--------|-----------------|----------|--------|----------|----------------|----------|--------|----------|
| | Bias | | Bredde | | Bias | | Bredde | |
| Sann | Post | Post/Pre | Post | Post/Pre | Post | Post/Pre | Post | Post/Pre |
| 10.355 | 0.021 | 0.976 | 1.033 | 0.999 | 0.123 | 6.034 | 1.800 | 1.100 |

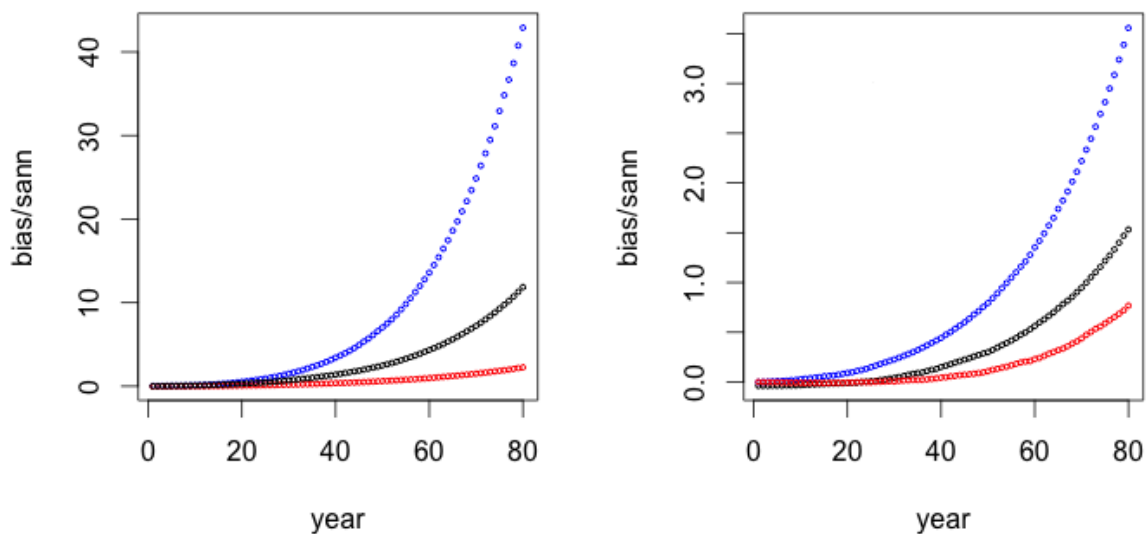


Figur 6: Π_{x0} , skjevhet (til venstre) og bredde på konfidensintervall og fremskrivningsintervall (i turkis) (til høyre), begge i prosent av forventningsverdien under modellen i tabell 1, for pre- (i rødt), post- (i svart) og uglattet (i blått) Lee-Carter i scenario 4.

Det kan også være interessant å se på langsiktige forpliktelser. Figur 7 viser utviklingen i skjevhet i dødeligheten for en 18-åring til venstre og en 60-åring til høyre. Forholdet mellom skjevhet i q_{xk} ved post- og pre-glatting for aldersgruppen 18-80 er omtrent det samme. For aldersgruppen 81-105 fører pre-glatting, særlig for de aller eldste, til større skjevhet i q_{xk} enn post-glatting, jamfør figur 3, som viser årsaken til at pre-glatting sammenliknet med post-glatting vil føre til underestimering (til høyre) etterfulgt av overestimering (til venstre) for denne aldersgruppen. Figur 7 antyder at topp- og bunnpunkt til skjevhet-kurvene til venstre i figur 6 vil skiftes nedover, og at skiftet vil være betydelig større ved post-glatting. For å bekrefte dette beregnes Π_{xk} for $x = 18, 60$ og $k = 60, 80$ ved både post- og pre-glatting, se tabell 5. Tabell 5 viser at etter 60 år er bunnpunktet ved post-glatting godt under det for pre-glatting, og etter 80 år er også toppunktene på nivå. Pre-glatting resulterer også på lang sikt i en reduksjon av risikoen på rundt 10% relativt til post-glatting.

Tabell 5: Langsiktige forpliktelser: Π_{xk} , $x = 18, 60$ og $k = 60, 80$. Skjevhet i prosent av forventningsverdi for scenario 4.

| $k = 60$ | | | | $k = 80$ | | | |
|----------|------|----------|------|----------|------|----------|------|
| $x = 18$ | | $x = 60$ | | $x = 18$ | | $x = 60$ | |
| Post | Pre | Post | Pre | Post | Pre | Post | Pre |
| -5.0 | -4.5 | -2.2 | -2.6 | -6.0 | -5.0 | -3.0 | -3.0 |



Figur 7: Skjevhet/forventningsverdi under modellen i tabell 1, ved bruk av uglattet (i blått), post- (i svart) og pre-glattet (i rødt) Lee-Carter i scenario 4, \hat{q}_{18k} (til venstre) og \hat{q}_{60k} (til høyre).

3.2 Glatting og parametrisering

Dess mer fleksibel modellen er jo bedre vil sannsynligvis tilpasningen til datasettet være, men desto større er også risikoen for å fange opp og fremskrive tilfeldige eller flyktige trender. Denne risikoen øker uansett med minkende størrelse på datagrunnlaget, men er økningens bratthet uavhengig av modellens parametriseringsgrad? Eller er det en sammenheng mellom modellens parametriseringsgrad og modellens evne til å håndtere tilfeldig variasjon?

3.2.1 Glatting av Lee-Carters modell

For å utforske sammenhengen mellom modellens parametriseringsgrad og modellens evne til å håndtere tilfeldig variasjon vil vi først sammenlikne en uglattet versjon av Lee-Carter modellen med den post-glattede versjonen fra avsnitt 3.1. Den post-glattede versjonen velges fremfor den pre-glattede slik at effekten av ett-trinns estimering holdes utenfor. Den uglattede versjonen er ikke-parametrisk og har $2n_x + 2$ parametre, mens den glattede er semi-parametrisk og har $3 + n_s + 2$ parametre og dermed en høyere parametriseringsgrad. Det vil ikke være overraskende om det i påfølgende eksperiment viser seg at post-glatting fører til mer stabile og regelmessige estimater av a_x og spesielt b_x , og at dette forplanter seg til fremskrevet dødelighet. Dette er et kjent fenomen, se avsnitt 2.2 med referanser. Det er derimot ikke, til min viten, gjort noe forsøk på å kvantifisere hvordan denne effek-

ten avhenger av størrelsen på datagrunnlaget. På samme måte som i avsnitt 3.1 er det utført et simuleringseksperiment der virkeligheten er antatt å følge Lee-Carter modellen i likning (1)-(4) med parameterverdier som i tabell 1. Semi-parametrisk bootstrapping er utført med uglattet Lee-Carter på samme måte som med post- og pre-glattet i avsnitt 3.1.

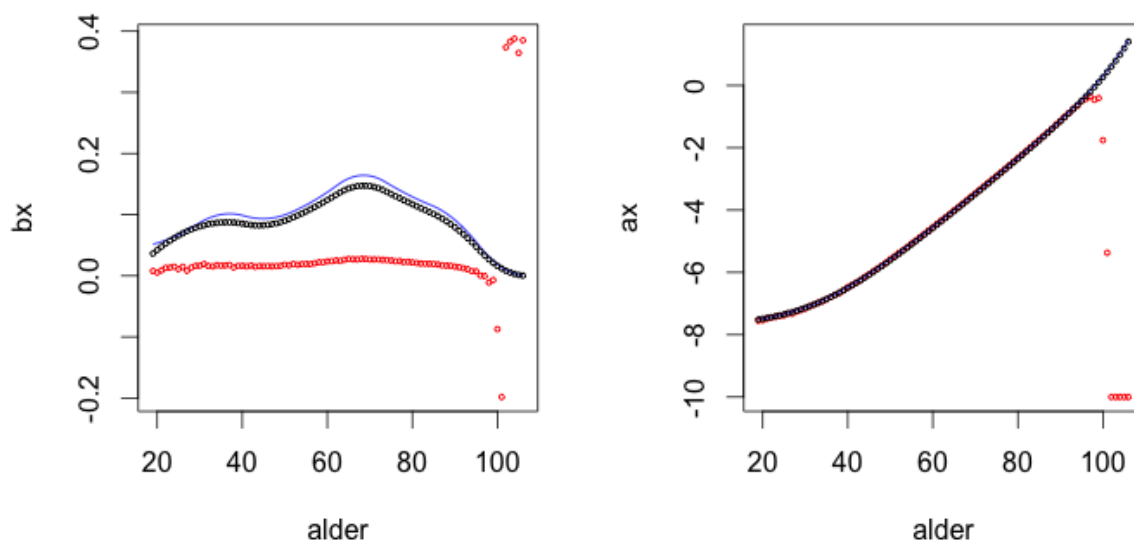
Tabell 6 viser at mangel på glatting resulterer i en grov overestimering, som øker ettersom datagrunnlaget begrenses, av begge de dynamiske parametrene ζ og σ . I scenario 4 vil mangel på glatting resultere i at usikkerheten i fremtidig dødelighet overdrives med en faktor på rundt 9. Post-glatting fører til mye mindre skjevhet og også mindre usikkerhet i parameterestimaterne, og effekten er større når datagrunnlaget begrenses og spesielt utpreget for scenario 3. At en uglattet Lee-Carter modell er ustabil med et datagrunnlag på 1.5 millioner pr år og 20 år med data, er ikke overraskende. Lee and Carter (1992) observerte at en baseperiode på 10-20 år resulterer i ustabile fremskrivninger selv for den amerikanske populasjonen. At post-glatting resulterer i en så betydelig forbedring, vitner kanskje om at god håndtering av korte baseperioder krever en mindre fleksibel modell.

Det fremkommer av tabell 7 og figur 8 at mangel på glatting ikke bare påvirker de dynamiske parametrene. Post-glatting av b_x øker stabiliteten på estimatene for alle scenarier. Venstre side av figur 8 viser graden av underestimering av b_x i scenario 4. Det ser ut til at underestimering av b_x "kompenseres" med overestimering av ζ og σ . Legg også merke til de store uregelmessighetene i de uglattede b_x estimatene for de eldste i figur 8.

Både post- og uglattet Lee-Carter underestimerer a_x i scenario 4, men post-glatting forhindrer at startdødeligheten modelleres unaturlig lav for de eldste aldersgruppene, se høyre side av figur 8. Denne effekten av glatting øker naturlig nok etter som datagrunnlaget begrenses.

Tabell 6: *Dynamiske parametre; Skjevhet (bias) ved bruk av uglattet Lee-Carter, omtrentlig forhold mellom skjevhet for uglattet og post-glattet Lee-Carter, og standardfeil (se) ved bruk av uglattet Lee-Carter.*

| | 1.5 mill, 50 år | | | 150 000, 50 år | | | 1.5 mill, 20 år | | | 150 000, 20 år | | |
|----------|-----------------|-------|--------|----------------|-------|--------|-----------------|-------|--------|----------------|-------|--------|
| | bias | /post | se | bias | /post | se | bias | /post | se | bias | /post | se |
| ζ | 0.0557 | 31 | 0.0461 | 0.4102 | 373 | 0.1357 | 0.4943 | >4900 | 0.2923 | 0.6862 | 89 | 0.4230 |
| σ | 0.1353 | 10 | 0.0928 | 1.2997 | 10 | 0.1449 | 1.0487 | 89 | 0.2129 | 2.1460 | 12 | 0.2311 |

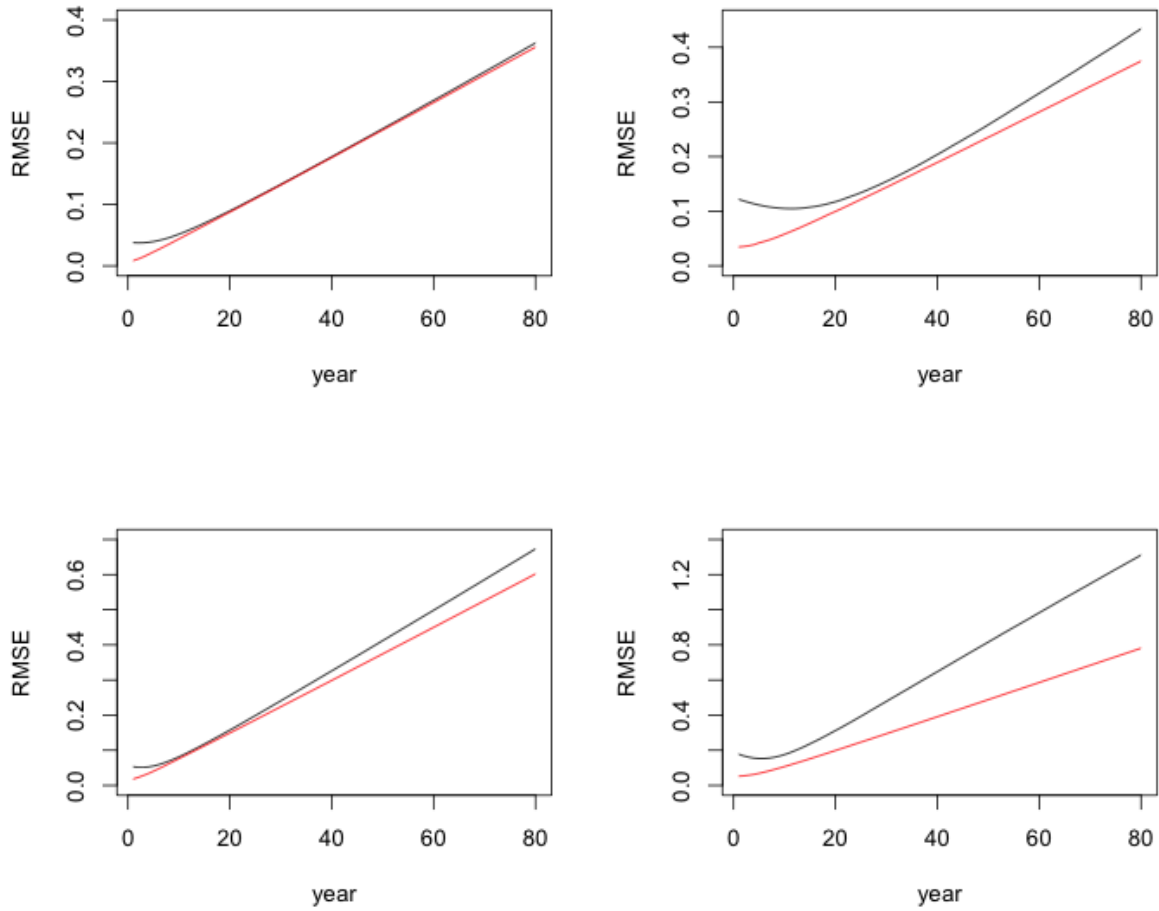


Figur 8: b_x for modellen i tabell 1 (i blått) og gjennomsnittlig b_x ved uglattet (i rødt) og post-glattet (i svart) Lee-Carter i scenario 4 (til venstre). a_x for modellen i tabell 1 (i blått) og gjennomsnittlig a_x ved uglattet (i rødt) og post-glattet (i svart) Lee-Carter i scenario 4 (til høyre).

Figur 9 viser hvordan den observerte fordelten med post-glatting forplanter seg til fremskrevet $\log(q_{xk})$, og effekten øker tydelig etter som datagrunnlaget minker og tilfeldig variasjon øker. Mangel på glatting fører som post-glatting til underestimering etterfulgt av overestimering av $\log(\hat{q}_k)$, se figur 5. Graden av overestimering er imidlertid mye mer alvorlig uten glatting. Som for post-glatting finnes forklaringen i skjevhet i parameterestimaten. For en kort fremskrivningshorisont er $\log(\hat{q}_{xk})$, som for post-glatting, preget av underestimering av a_x . Etter som fremskrivningshorisonten øker preges $\log(\hat{q}_{xk})$ av samspillet mellom b_x og t_k . Overestimeringen av ζ og σ med en faktor på henholdsvis 5 og 8 kompenserer tydeligvis ikke helt for underestimeringen av $\{b_x\}$ med en faktor på mellom 4 og 19 (sett bort i fra de eldste aldersgruppene).

Tabell 7: $RMSE_b$ for uglattet Lee-Carter og omtrentlig forhold mellom $RMSE_b$ for uglattet og post-glattet Lee-Carter.

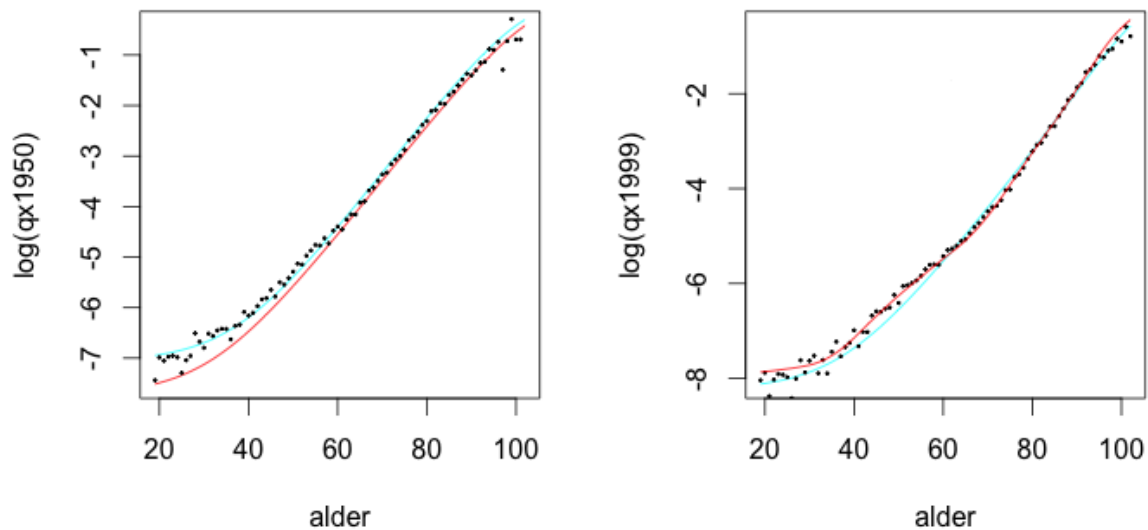
| 1.5 mill, 50 år | | 150 000, 50 år | | 1.5 mill, 20 år | | 150 000, 20 år | |
|-----------------|----------|----------------|----------|-----------------|----------|----------------|----------|
| Raw | Raw/Post | Raw | Raw/Post | Raw | Raw/Post | Raw | Raw/Post |
| 0.0642 | 19.6272 | 0.1339 | 6.7294 | 0.1260 | 10.9097 | 0.1335 | 4.6750 |



Figur 9: $RMSE_{\log(\hat{q}_k)}$ for uglattet (i svart) og post-glattet (i rødt) Lee-Carter, scenario 1 (øverst til venstre), scenario 2 (øverst til høyre), scenario 3 (nederst til venstre) og scenario 4 (nederst til høyre).

Som i avsnitt 3.1.1 observeres samme trend for $RMSE_{\hat{e}_{18k}}$, som for $RMSE_{\log(\hat{q}_k)}$. Effekten av post-glutting akselererer med minkende datagrunnlag og ettersom fremskrivningshorisonten øker. Høyre side av figur 5 viser, i samsvar med venstre side, at mangel på glutting resulterer i en kraftig underestimering av \hat{e}_{18k} , og etter 80 år underestimeres \hat{e}_{18k} med mer enn 6 år. Glutting resulterer igjen i mye mindre skjevhet.

Det ser ut til at den mindre fleksible post-glattede modellen er overlegen i håndtering av tilfeldig variasjon. For å se på potensielle økonomiske konsekvensene beregnes det, som i avsnitt 3.1.2, dynamiske engangspremier, se figur 6. Ved å post-glutte a_x og b_x reduseres bredden på konfidensintervallet rundt Π_{x0} med et gjennomsnitt på rundt 7%, se høyre side av figur 6. Fra venstre side av figur 6 fremkommer det, ikke uventet, at mangel på glutting også resulterer i en underestimering av engangspremien for alle aldre under 61, og for de yngste aldersgruppene med opp til 4% av forventningsverdien. Mangel på glutting kan dermed ha store økonomiske konsekvenser for et livsforsikringsselskap. Forholdet mel-



Figur 10: $\log(q_{xk})$, sveitsiske kvinner (i svart), tilpasning under den parametriske modellen (i turkis) i tabell 8 og under den pre-glattede versjonen av Lee-Carter modellen i tabell 1 (i rødt), $k = 1950$ (til venstre) og $k = 1999$ (til høyre).

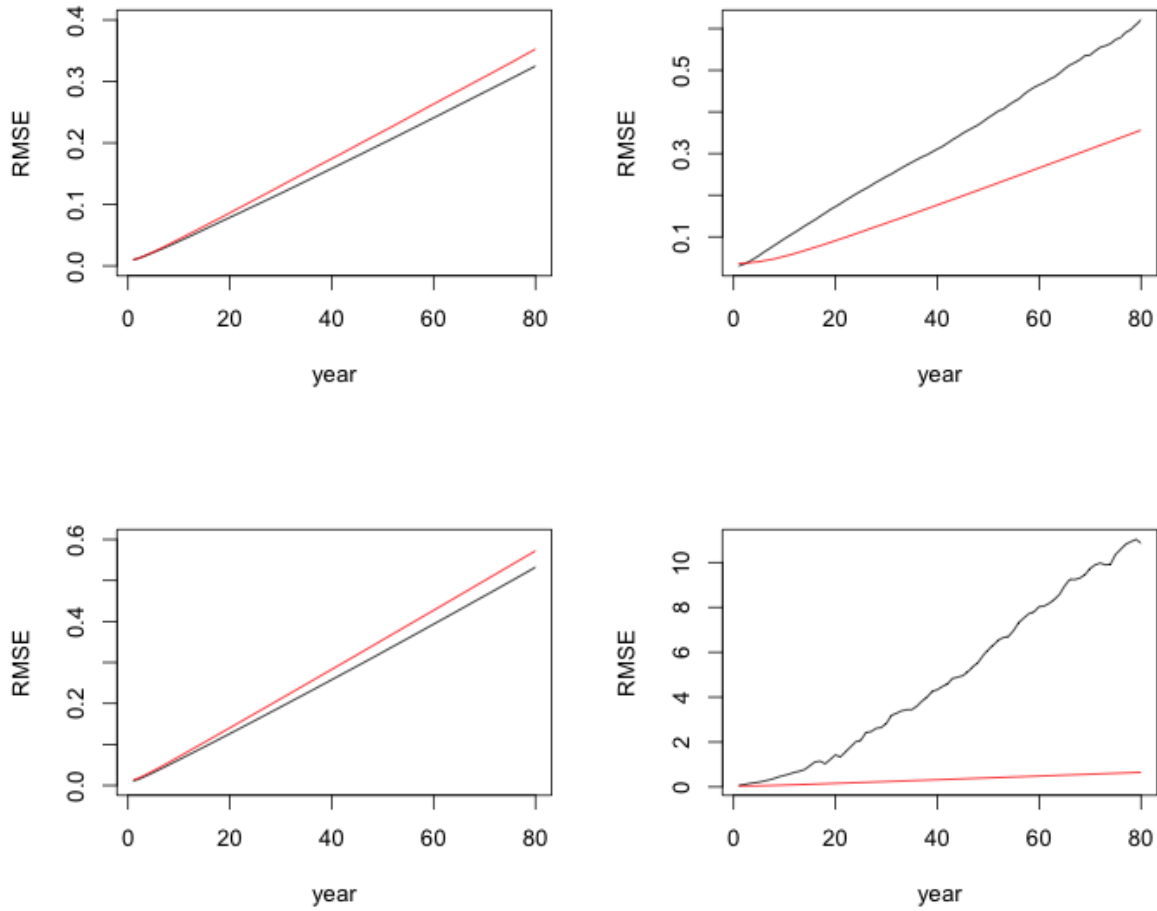
lom skjevhet i q_{xk} ved post- og uglattet Lee-Carter i figur 7 er omtrent det samme for alle aldersgrupper, og de økonomiske konsekvensene vil dermed bare øke med fremskrivningshorisonten.

3.2.2 Pre-glattet Lee-Carter mot parametrisk

Lee-Carter modellen er som tidligere nevnt motivert ut fra et meget stort land, og har senere vist seg å være potensielt ustabil når datagrunnlaget begrenses, se avsnitt 2.2 med referanser. Avsnitt 3.2.1 viser at stabiliteten til Lee-Carter type modeller kan forbedres ved å begrense modellens fleksibilitet. Et naturlig spørsmål er derfor om en parametrisk og mindre fleksibel modell vil håndtere tilfeldig variasjon bedre enn glattede Lee-Carter type modeller. For å belyse dette spørsmålet vil det her bli gjort et forsøk på å sammenlikne stabiliteten til pre-glattet Lee-Carter med den parametriske modellen i avsnitt 2.3.

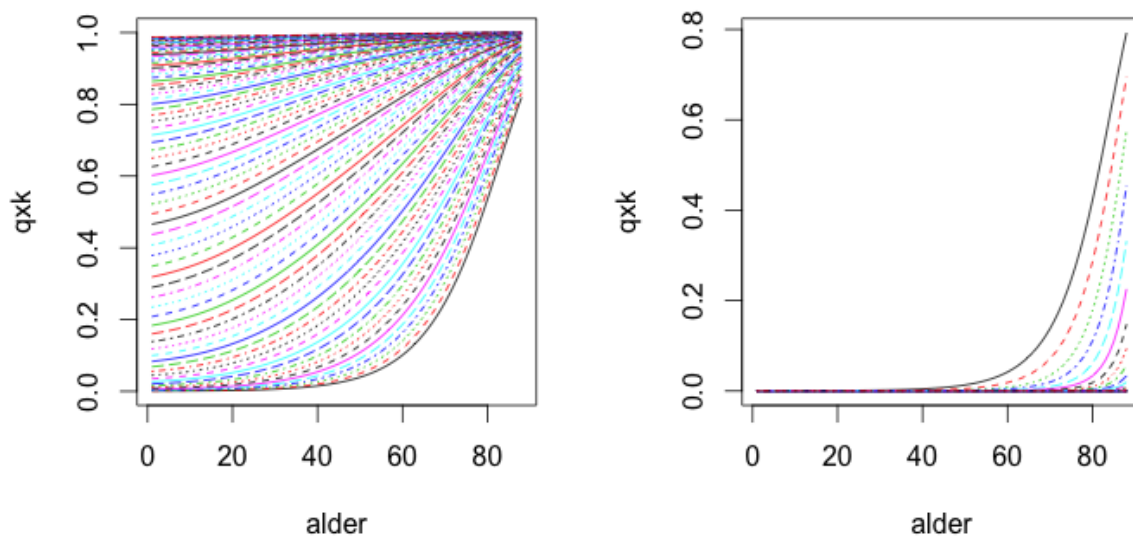
Tabell 8: Den parametriske modellen (med begrensninger) tilpasset sveitsiske kvinner.

| ζ | σ | γ | $10^3\theta_0$ | $10^3\theta_1$ | θ_2 |
|------------------|-------------|-------------------|------------------|------------------|-------------------|
| 0.0146(-0.5,0.5) | 0.0184(0,1) | 1.6587(0.0001,10) | 0.8530(0.0003,6) | 0.0980(0.0003,6) | 0.1134(0.006,0.2) |



Figur 11: $RMSE_{\log(\hat{q}_k)}$ for den parametriske modellen (i svart) og pre-glattet Lee-Carter (i rødt), scenario 1 (øverst til venstre), scenario 2 (øverst til høyre), scenario 3 (nederst til venstre) og scenario 4 (nederst til høyre).

For pre-glattet Lee-Carter vil resultatene i avsnitt 3.1 anvendes. For den parametriske modellen er det utført samme simuleringseksperiment som for Lee-Carter modellen, men virkeligheten er nå antatt å følge den alternative parametriske modellen. Parameterverdiene er som for Lee-Carter basert på populasjonen av sveitsiske kvinner mellom 18-105 år for perioden 1950-1999, se tabell 8. Det er, som vist i tabell 8, satt begrensninger på alle parametrene, og antall parametre er 6 mot $3 + n_s + 2$ i glattet Lee-Carter. Figur 10 viser overensstemmelse mellom virkelighet og tilpasning under den parametriske modellen i tabell 8 og Lee-Carter modellen i tabell 1 for første og siste baseår.

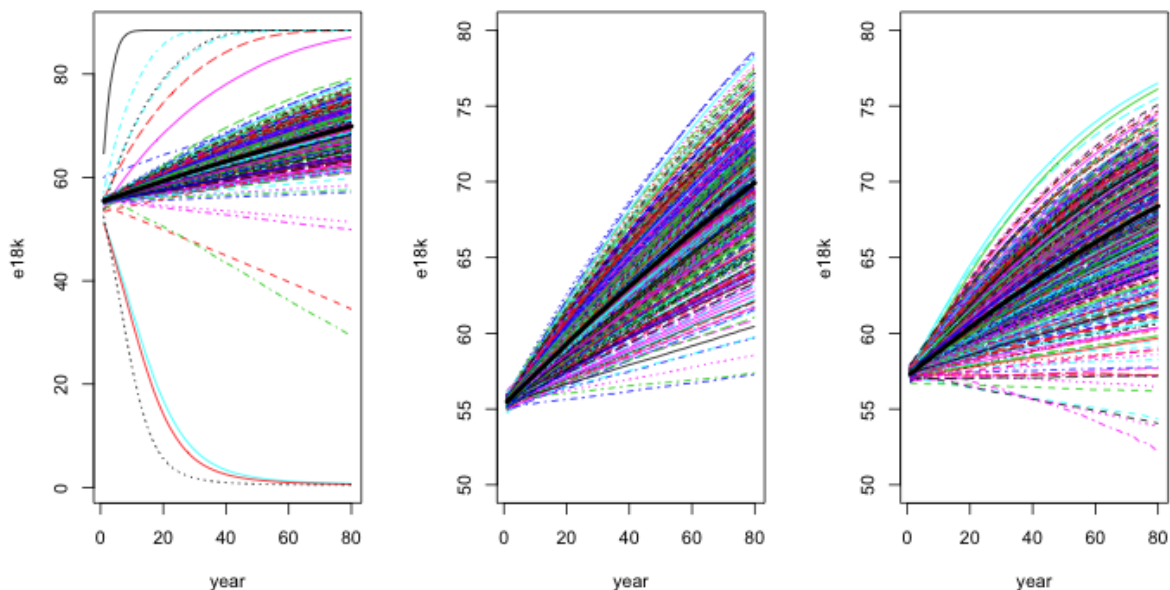


Figur 12: *Fremskrevet dødelighet for alle aldre 80 år frem i tid, hver graf viser dødeligheten for alle aldre i et gitt år. Til venstre: dødeligheten går mot 1 for alle aldre, $\zeta = -0.0858$, $\sigma = 0.0348$ og $\gamma = 9.9999$. Til høyre: dødeligheten går mot 0 for alle aldre, $\zeta = 0.4221$, $\sigma = 0.3547$ og $\gamma = 10.0000$.*

Tabell 9 viser skjevhet og standardfeil til de dynamiske parametrene under den parametriske modellen for de fire scenariene. Grunnet modellens formulering, de dynamiske egenskapene avhenger av γ , kan det ikke foretas en direkte sammenlikning mellom tabell 9 og 2, men tabellene kan vitne om at den parametriske modellen er relativt ustabil når antall personer pr år begrenses (scenario 2 og 4). Figur 11 antyder at den parametriske modellen håndterer en kort baseperiode bedre enn den pre-glattede Lee-Carter modellen så lenge antall personer pr år ikke begrenses. Når antall personer pr år begrenses, når variasjonen i dødelighet er stor fra år til år, ser det imidlertid ut til at den parametriske modellen feiler. Bootstrap-realiseringene for scenario 2 og 4 analyseres nærmere, og det viser seg at en del av bootstrap-modellene resulterer i situasjoner som i figur 12 og venstre side av figur 13.

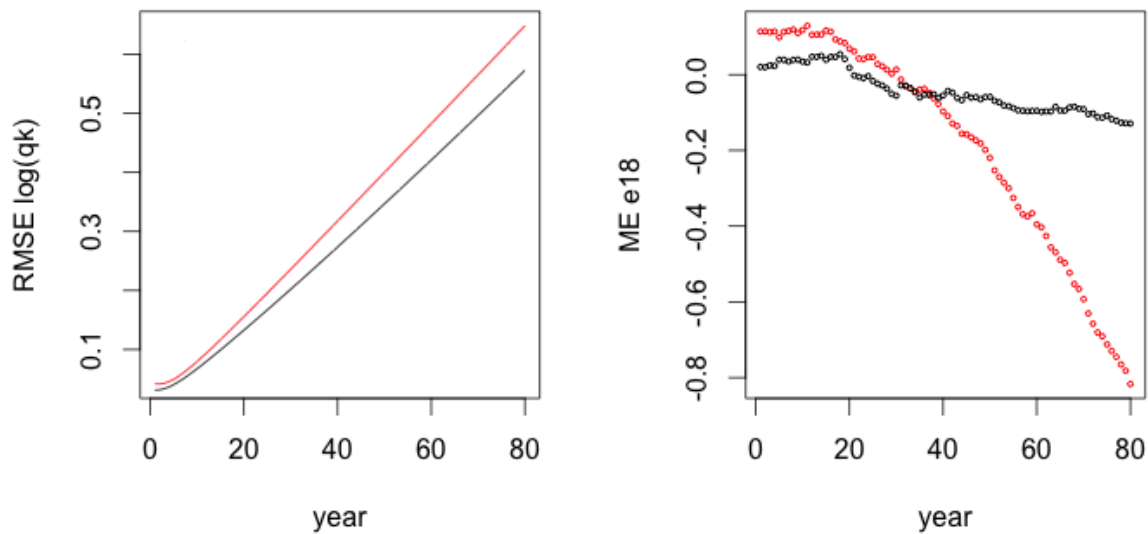
Tabell 9: *Skjevhet (bias) og standardfeil (se) for de dynamiske parametrene under den parametriske modellen.*

| | 1.5 mill, 50 år | | 150 000, 50 år | | 1.5 mill, 20 år | | 150 000, 20 år | |
|----------|-----------------|--------|----------------|--------|-----------------|--------|----------------|--------|
| | bias | se | bias | se | bias | se | bias | se |
| ζ | -0.0002 | 0.0026 | -0.0013 | 0.0035 | -0.0003 | 0.0043 | 0.0931 | 0.1890 |
| σ | -0.0005 | 0.0021 | -0.0024 | 0.0048 | -0.0012 | 0.0036 | 0.097 | 0.2477 |
| γ | 0.0221 | 0.0377 | 0.2300 | 0.5390 | 0.0535 | 0.1828 | 0.2170 | 1.7776 |



Figur 13: *Bootstrap-realiseringer for \hat{e}_{18k} i scenario 4: den parametriske modellen i tabell 8 (til venstre), øvre begrensninger på dynamiske parametre hevet: $\gamma = 100\,000$, $\zeta = 1$ og $\sigma = 2$ (midten), pre-glattet versjon av Lee-Carter modellen i tabell 1 (til høyre), svart tykk linje representerer virkeligheten under de gitte modellene.*

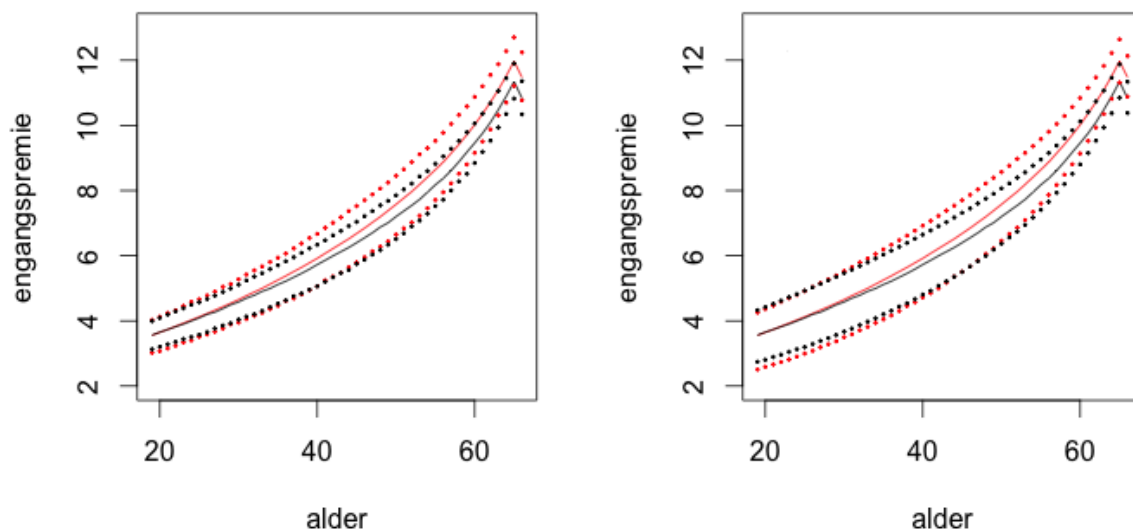
At dødeligheten skal gå til 0 eller 1 for alle aldre i løpet av 80 år er ekstremt urealistisk. Scenariene i figur 12 er uten unntak assosiert med at et av de dynamiske parametrene ζ , σ eller som regel γ når sin begrensning. Basert på dette kan det mistenkes at begrensningen på γ , og eventuelt ζ og σ , er satt for lav. Innen dødelighetsmodellering er utfordringer assosiert med flate likelihoodoverflater ikke et ukjent fenomen, se for eksempel Cairns et al. (2011a) for et praktisk eksempel. Det kan tenkes at noen av bootstrap-realiseringene gir opphav til spesielt flate likelihoodoverflater med multiple alternative maksima, og at modellens begrensninger resulterer i at et ugunstig alternativ velges. Det observeres at algoritmen generelt bruker flere iterasjoner på å konvergere når antall personer pr år begrenses, noe som kan antyde flate likelihoodoverflater. Det eksperimenteres med alternative øvre begrensninger på γ , ζ og σ . Med øvre begrensning på $\gamma = 100\,000$, $\zeta = 1$ og $\sigma = 2$ forandres situasjonen totalt. Midtre og høyre del av figur 13 indikerer at den parametriske modellen nå takler tilfeldig variasjon vel så godt som en pre-glattet Lee-Carter modell. Bootstrap-modellene, som midtre del av figur 13 er basert på, antyder også at problemene stammet fra flate likelihoodoverflater. Med de nye parameterbegrensningene er for eksempel maksimal verdi på γ , blant $N = 500$ bootstrap-modeller, under 7. Tilsvarende resultater oppnås også for scenario 2.



Figur 14: $RMSE_{\log(\hat{q}_k)}$ (til venstre) og $ME_{\hat{e}_{18k}}$ (til høyre) i scenario 4, for den pre-glattet versjonen av Lee-Carter modellen (i rødt) og for den parametriske modellen med nye parameterbegrensninger (i svart).

For å utelukke at denne løsningen ikke bare er et heldig engangstilfelle, for modellen basert på sveitsiske kvinner, analyseres også stabiliteten til den parametriske modellen basert på norske menn, samme aldersgruppe og periode som for sveitsiske kvinner, (kun hovedfunn rapporteres her). Samme situasjon som for sveitsiske kvinner observeres. Med parameterbegrensninger som i tabell 8 observeres flere tilfeller av scenariene i figur 12. Dette korrigeres igjen ved å øke de øvre parameterbegrensningene på ζ og σ til henholdsvis 1 og 2, men i dette tilfellet må øvre begrensning på γ heves til 1 000 000. En grundigere analyse er nødvendig for å fastsette modellens optimale parameterbegrensninger (gjeldende uansett datasett), men for denne oppgavens formål er vi nære nok.

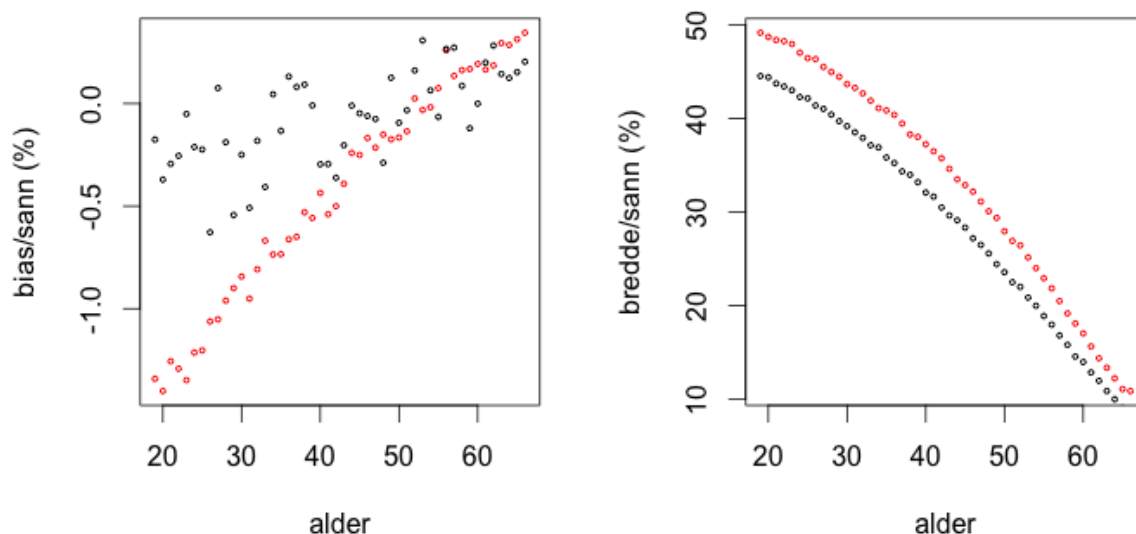
Venstre side av figur 14 viser at bruk av den parametriske modellen i scenario 4 nå resulterer i mer nøyaktige estimer av $\log(\hat{q}_{xk})$ enn en pre-glatt Lee-Carter. Dette gjelder også, men i en noe mindre grad, for de tre andre scenariene (ikke vist her). Samme trend observeres for $RMSE_{\hat{e}_{18k}}$. Fra høyre side av figur 14 og midtre og høyre del av figur 13 fremkommer det at skjevhet i \hat{e}_{18k} i scenario 4 under den parametriske modellen er tilnærmet neglisjerbar relativt til skjevhet under pre-glatt Lee-Carter. Det samme gjelder for scenario 2 og 3, mens det i scenario 1 er lite skjevhet uansett valg av modell. Det er kjent at Lee-Carter modellen kan ha en tendens til å underestimere forventet gjenværende levetid for virkelige data, se for eksempel Lee and Miller (2001), Booth et al. (2005) eller Booth et al. (2006), men i dette tilfellet er virkeligheten antatt å følge Lee-Carter modellen. At scenario 1 er tilnærmet skjevhetsfritt mens det observeres en økende grad av negativ skjevhet i scenario 2, 3 og 4 vitner dermed bare om at Lee-Carter modellen er underlegen den parametriske modellen i håndtering av tilfeldig variasjon.



Figur 15: Π_{x0} under Lee-Carter modellen i tabell 1 (i rødt) og den parametriske modellen i tabell 8 (i svart), med fremskrivningsintervall (til venstre) og konfidensintervall med ett-trinns estimering for scenario 4 (til høyre).

3.2.3 Pre-glattet Lee-Carter mot parametrisk: Økonomisk

For å kartlegge potensielle økonomiske konsekvenser beregnes det, som tidligere, dynamiske engangspremier. Fra venstre side av figur 15 fremkommer det at den virkelige verdien på Π_{x0} er noe lavere under den parametriske modellen. Forskjellen øker med alderen, fra under 1% for aldersgruppen under 25 til opp mot 6% for aldersgruppen over 40. Figur 10 antyder, og nærmere analyse bekrefter, at grunnen er at Lee-Carter underestimerer dødeligheten for hele den aktuelle aldersgruppen i de første baseårene. Dette har, naturlig nok, mindre konsekvens for engangspremiene for de yngste, ettersom disse i større grad er avhengig av fremtidig dødelighet. I samsvar med trenden for $ME\hat{e}_{18k}$, se høyre side av figur 14, viser venstre side av figur 16 at skjevhet i dynamisk engangspremie for aldersgruppen 18-65 under den parametriske modellen er ubetydelig sammenliknet med den pre-glattede versjonen av Lee-Carter modellen. Det samme gjelder for langsiktige forpliktelser. Skjevhet i $\Pi_{18,60}$ og $\Pi_{60,60}$ er på henholdsvis -1.5% og -0.5% av forventningsverdien under den parametriske modellen, mot -4.5% og -2.6% under pre-glattet Lee-Carter.



Figur 16: Π_{x0} , $x = 18, \dots, 65$: Skjevhet (til venstre) og bredde på konfidensintervall (til høyre) for pre-glattet Lee-Carter (i rødt) og den parametriske modellen (i svart) i prosent av forventningsverdi under henholdsvis Lee-Carter modellen i tabell 1 og den parametriske modellen i tabell 8, scenario 4 .

Videre resulterer bruk av den parametriske modellen i en reduksjon av bredden på konfidensintervallene rundt Π_{x0} , med en gjennomsnittlig verdi på rundt 20% sammenliknet med pre-glattet Lee-Carter, se høyre side av figur 16 og 15. En 20% risikoreduksjon er såpass at det kan ha betydning for prisen på en enkelt livsvarig annuitet. Høyre side av figur 15 og venstre side av figur 16 indikerer også at konfidensintervallet under Lee-Carter modellen ikke er symmetrisk rundt punktestimatet. Det viser seg at under Lee-Carter modellen er nedre del av konfidensintervallet i gjennomsnitt rundt 20% bredere enn øvre del, mens konfidensintervallet er tilnærmet symmetrisk under den parametriske modellen. I verste fall (nedre grense av konfidensintervallet) står selgeren av en portefølje av livsvarige annuiteter, bestående av rundt 130 000 sveitsiske kvinner mellom 18-65, overfor et tap som er anslagsvis 30% høyere hvis pre-glattet Lee-Carter anvendes for prising istedenfor den parametriske modellen

$$\frac{\sum_{x=18}^{65} (\hat{\Pi}_{x0} - \hat{\Pi}_{x0}^{0.025*N})_{lee} N_{x0}}{\sum_{x=18}^{65} (\hat{\Pi}_{x0} - \hat{\Pi}_{x0}^{0.025*N})_{par} N_{x0}} = 1.32.$$

Dette er selvsagt bare indikasjoner ettersom vi ikke tar hensyn til modellfeil, men gir likevel en antydning om at modellens parametriseringsgrad kan ha økonomiske konsekvenser.

4 Historisk analyse av noen land

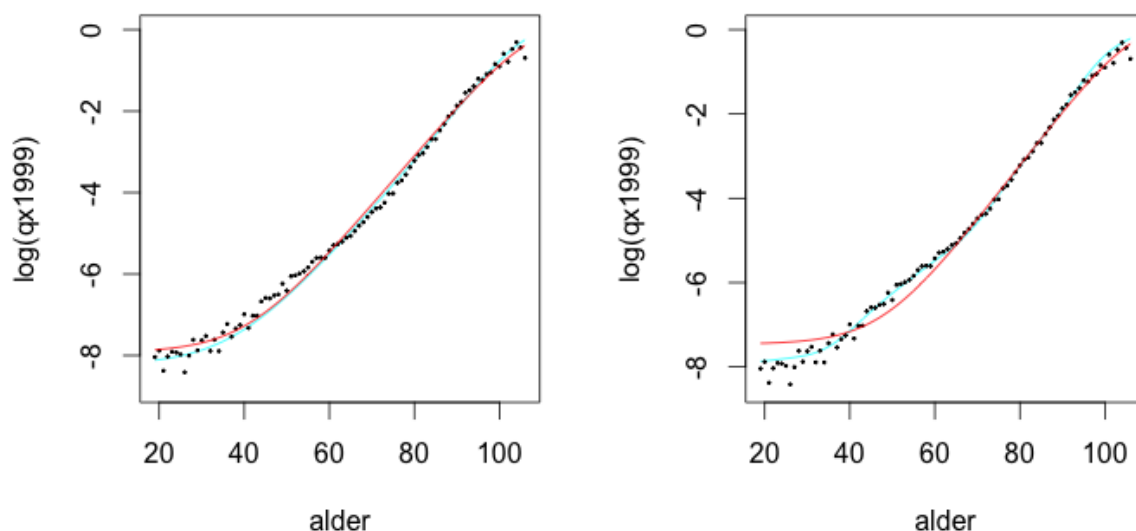
4.1 Innledning

En god stokastisk dødelighetsmodell bør være konsistent med historiske data, men til hvilken grad? Hvis antall tilgjengelige år med data er begrenset, kan en “god” modell fange opp flyktige trender, og den langsiktige dynamikken under modellen vil ikke nødvendigvis være biologisk rimelig. Hvis det i tillegg er et begrenset antall personer pr år, kan en “god” modell også fange opp tilfeldige trender. Avsnitt 3 antyder at Lee-Carter modellens evne til å håndtere tilfeldig variasjon kan forbedres ved å begrense modellens fleksibilitet, og at denne begrensningen (glattingen) bør foretas under og ikke etter parameterestimeringen. Videre antydes det at Lee-Carter type modeller kanskje bør erstattes med mindre fleksible parametriske modeller når datagrunnlaget er begrenset.

Analysen i avsnitt 3 er basert på numeriske eksperimenter der virkeligheten er antatt å følge Lee-Carter modellen eller det parametriske alternativet. Men virkeligheten følger ingen kjent modell. Det vil overraske dersom glatting ikke begrenser Lee-Carter modellens evne til å tilpasse seg historiske data, se for eksempel Currie (2013). Formålet med å begrense modellens fleksibilitet er nettopp å forhindre at modellen fanger opp tilfeldige eller flyktige trender slik at modellens langsiktige dynamikk blir mer biologisk rimelig. Men vil glatting eller parametrisering ha denne effekten når modellene tilpasses virkelige datasett, eller vil det bare føre til en generelt dårligere tilpasning? Og hvilken effekt har ett-trinns estimering på modellens evne til å tilpasse seg historien?

For å utforske dette nærmere er det foretatt en ex-post analyse av tilpasnings- og fremskrivningsevne. I tråd med oppgavens overordnede fokus har vi tatt for oss datasett av varierende størrelse, populasjonene av kvinner og menn hver for seg i Frankrike, Sveits og Island. Disse tre representerer her store, medium og små land med et omtrentlig gjennomsnitt på henholdsvis 18 millioner, 2 millioner og 70 000 kvinner og menn pr år for aldersgruppen 18-105 og perioden 1950-1999. Datasettene er basert på data fra Human Mortality Database (HMD), se appendiks C for mer informasjon om HMD og initiell databehandling.

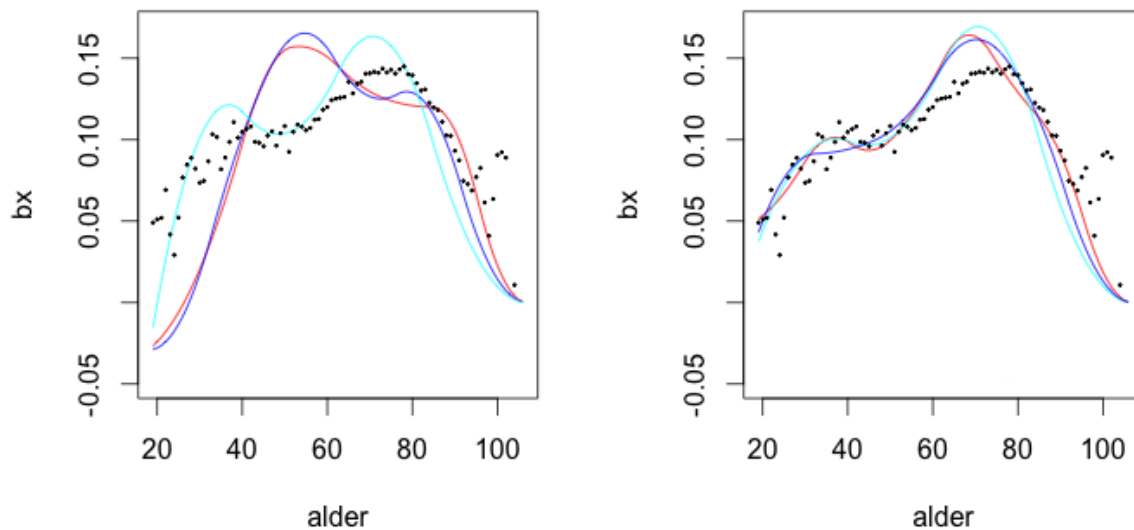
I samsvar med Lee and Miller (2001) settes 1950 som tidligste mulige startpunkt for tilpasningsperioden, se avsnitt 2.2. For å evaluere fremskrivningsevne er det ønskelig med en relativt lang fremskrivningshorisont, ettersom dynamisk forventet gjenværende levetid og engangspremie for eksempelvis en 18-åring avhenger av $q_{(18+i)(k+i)}$ for $i = 0, \dots, (n_x - 1)$ i fremskrivningsår k . For å best mulig kunne se resultatene i sammenheng med analysen i avsnitt 3 anvendes det også her tilpasningsperioder på 50 og 20 år, 1950-1999 og 1970-1989, med tilhørende maksimal fremskrivningshorisont på henholdsvis $n_l = 11$ og $n_l = 21$ år (ettersom siste tilgjengelige år med data er 2010), heretter referert til som scenario 1 og 2. Her kunne vi i stedet valgt å bruke for eksempel 1950-1969 som 20 års tilpasningsperiode og dermed hatt en maksimal fremskrivningshorisont på 41 år. Men ettersom det er ønskelig at 20-års perioden skal være noenlunde representativ for 50-års perioden, og også at start- og sluttpunktet på tilpasningsperiodene ikke sammenfaller, faller valget, litt tilfeldig, på 1970-1989. Vi kunne også valgt å følge Booth et al. (2002), valgt tilpasningsperioden som best oppfyller antakelsen om lineær t_k og konstant b_x , men det



Figur 17: $\log(q_{xk})$, sveitsiske kvinner (i svart), tilpasning i siste baseår, 1999, med opprinnelig (i turkis) og reversert (i rødt) datasett under den parametriske modellen (til venstre) og Lee-Carter modellen med $ns=8$ (til høyre).

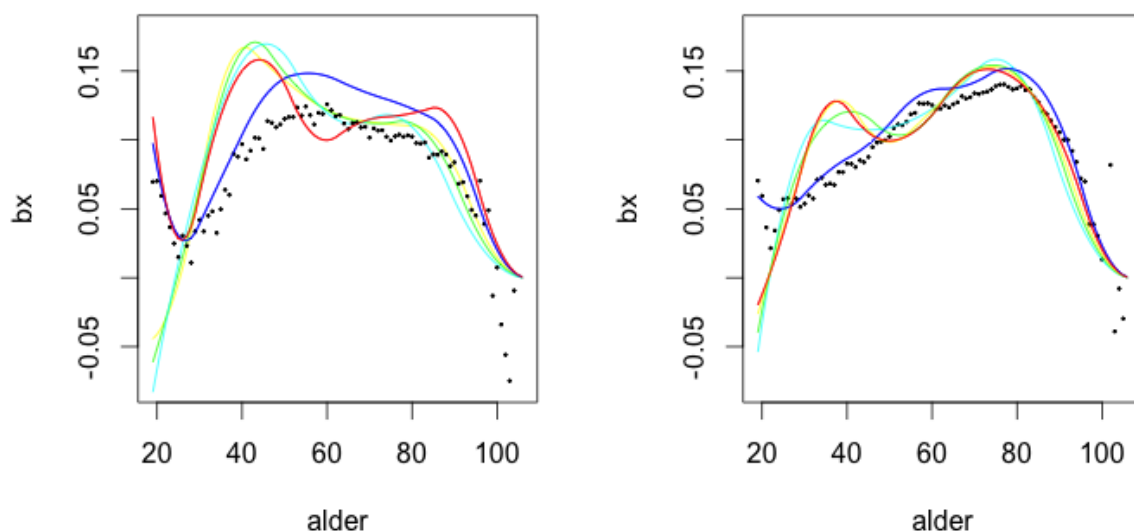
anses som heller urealistisk, fra et livsforsikringsperspektiv, at den tilgjengelige datamengden er av en størrelse som tillater valg av denne typen. Det er tvert i mot ønskelig å sammenlikne modellenes tilpasnings- og fremskrivningsevne på datasett som ikke er nøye utvalgt, og som dermed representerer en mer reell situasjon hvor det i mange tilfeller kan være ønskelig å anvende all tilgjengelig data.

For fremskrivningsformål kan det være hensiktsmessig å foreta estimeringen av modellens parametre slik at $t_{n_k} = 0$, se avsnitt 2.2 og 2.6. Poenget er å oppnå god tilpasning i siste baseår, slik at overgangen mellom faktisk dødelighet i siste tilpasningsår og fremskrevet dødelighet blir så glatt som mulig. For ett-trinns estimeringsmetoden i avsnitt 2.4 eksperimenteres det derfor med både opprinnelige og reverserte datasett. Det ser imidlertid ikke ut til at tilpasningen i siste baseår nødvendigvis forbedres ved å snu datasettet, se figur 17 for et eksempel. Det viser seg at modellenes dynamikk under ett-trinns estimering er sensitiv til retningen på datasettet. Konsekvensene er mer alvorlige for Lee-Carter modellen hvor reversering av datasett kan føre til reversering av dødelighetsutvikling (negativ b_x) for noen aldersgrupper, se høyre side av figur 17 og figur 18 for et eksempel.



Figur 18: *Pre-glattet Lee-Carter, b_x sveitsiske kvinner, 18-105 år og tilpasningsperiode 1950-1999, med reversert datasett (til venstre) og opprinnelig datasett (til høyre). (uglattet i svart, $n_s = 4$ i turkis, $n_s = 6$ i blått og $n_s = 8$ i rødt).*

Ved bruk av post-glatting er Lee-Carter modellens dynamiske egenskaper tilnærmet upåvirket av reversering av datasettet. Dette kan tyde på at ett-trinns estimeringsmetoden i avsnitt 2.4 genererer flatere likelihoodoverflater enn den tradisjonelle to-trinns metoden. En mulig indikasjon på dette så vi også i avsnitt 3.2. Figur 18 støtter opp om denne hypotesen ettersom det kan se ut til at to vidt forskjellige maksima oppnås ved opprinnelig og reversert datasett for $n_s = 6$ og 8 , mens de tilsvarende maksima med $n_s = 4$ ser ut til å være i nærheten av hverandre. Det samme fenomenet observeres også for blant annet franske menn, samme aldersgruppe og periode. For å unngå unødige komplikasjoner er modellene, for sammenlikningsformål, derfor tilpasset opprinnelige datasett og fremskrivningene begynner i $t_{n_k} = \tilde{t}_{n_k}$ (som i blant annet Booth et al. (2002)).



Figur 19: Glatting av b_x , uglattet i svart stiplet, post- og pre-glatting med $n_s = 8$ i henholdsvis blått og rødt, pre-glatting med $n_s = 5, 6$ og 7 i turkist, grønt og gult, sveitsiske menn (til venstre) og franske menn (til høyre).

Glatting av b_x er i ex-post analysen foretatt med $n_s = 8$ splinesegmenter. Dette fordi det viser seg at spline-funksjonen skifter form fra $n_s < 8$ til $n_s \geq 8$ når ett-trinns estimering anvendes for å tilpasse Lee-Carter modellen til sveitsiske menn mellom 18-105 år i perioden 1950-1999, se venstre side av figur 19. Bruk av $n_s \geq 8$ er i dette tilfelle gunstig, ettersom det da unngås at modellen forespeiler en økning i dødeligheten for de yngste, til tross for at trenden i datasettet sier det motsatte. Det samme gjelder imidlertid ikke når ett-trinns estimering anvendes på franske menn, samme aldersgruppe og tilpasningsperiode, der splinefunksjonens form er uforandret for $n_s = 4, \dots, 10$, se høyre side av figur 19. Begge tilfellene vitner igjen om at ett-trinns tilpasning av denne versjonen av Lee-Carter modellen kan være assosiert med flate likelihoodoverflater. Ettersom glattingsrutinen er konstruert slik at b_x uansett tvinges mot null for de eldste aldersgruppene, se avsnitt 2.2, er kanskje en mulig strategi, for å unngå situasjonen på høyre side av figur 19, å øke tettheten av vendepunkter rundt de yngre aldersgruppene. Hvis dette ikke er tilstrekkelig, er kanskje løsningen, som i Zhao (2012), en mer komplisert glattingsrutine, som muliggjør for eksempel valg av optimale vendepunkter og enderestriksjoner. Dette vil imidlertid ikke utforskes nærmere her.

For å få et overblikk over modellenes tilpasnings- og fremskrivningsevne beregnes RMSE fordelt på tilpasnings- og fremskrivningsperiode pr land og kjønn for logaritmen til dødeligheten,

$$RMSE_{\log(\dot{q})}^{til} = \sqrt{\frac{1}{n_x(n_k + 1)} \sum_{k=0}^{n_k} \sum_{x=1}^{n_x} (\log(\tilde{q}_{xk}) - \log(\dot{q}_{xk}))^2}$$

for hver av de to tilpasningsperiodene (1950-99: scenario 1 og 1970-89: scenario 2), og

$$RMSE_{\log(\dot{q})}^{frem} = \sqrt{\frac{1}{n_x n_l} \sum_{l=1}^{n_l} \sum_{x=1}^{n_x} (\log(\hat{q}_{xk}) - \log(\dot{q}_{xk}))^2}$$

for hver av de to tilhørende fremskrivningsperiodene (2000-2010: scenario 1 og 1990-2010: scenario 2), se tabell 10. Det kan virke unaturlig at $RMSE_{\log(\dot{q})}^{til}$ i noen tilfeller er høyere enn $RMSE_{\log(\dot{q})}^{frem}$. Dette grunner i store forskjeller mellom faktisk og tilpasset dødelighet for de eldste aldersgruppene (som amplifiseres på log-skala), og gjelder spesielt de første baseårene, der det er få eller ingen personer i disse aldersgruppene, se for eksempel figur 21. Tilsvarende størrelser beregnes også for statistisk e_{18} , se tabell 11. Statistisk e_{18} velges som aggregert mål ettersom det baseres på dødeligheten i alle aktuelle aldre, ikke er vektet (som Π_{18}), og ettersom relevante dynamiske størrelser naturlig nok ikke er tilgjengelige i en ex-post analyse.

Tabell 10: $RMSE_{\log(\dot{q})}$ fordelt på tilpasnings- og fremskrivningsperiode pr scenario, land og kjønn, under pre-, post- og uglattet Lee-Carter og den parametriske modellen, (Franske menn 1 er franske menn i scenario 1).

| | Tilpasning | | | | Fremskrivning | | | |
|----------------------|------------|--------|-------|-------|---------------|--------|-------|-------|
| | raw | post | pre | par | raw | post | pre | par |
| Franske menn 1 | 0.560 | 0.582 | 0.578 | 0.581 | 0.212 | 0.216 | 0.223 | 0.152 |
| Franske kvinner 1 | 0.260 | 0.296 | 0.272 | 0.302 | 0.134 | 0.212 | 0.175 | 0.263 |
| Sveitsiske menn 1 | 1.102 | 1.082* | 1.154 | 1.154 | 0.592 | 0.584* | 0.576 | 0.558 |
| Sveitsiske kvinner 1 | 0.702 | 0.742 | 0.720 | 0.731 | 0.236 | 0.347 | 0.252 | 0.214 |
| Islandske menn 1 | 1.624 | 1.313 | 1.312 | 1.322 | 1.729 | 1.728 | 1.681 | 1.719 |
| Islandske kvinner 1 | 1.510 | 1.524 | 1.523 | 1.540 | 2.068 | 2.129 | 2.160 | 2.139 |
| Franske menn 2 | 0.282 | 0.311 | 0.297 | 0.306 | 0.206 | 0.235 | 0.244 | 0.146 |
| Franske kvinner 2 | 0.075 | 0.138 | 0.103 | 0.179 | 0.170 | 0.218 | 0.239 | 0.289 |
| Sveitsiske menn 2 | 0.977 | 1.040 | 1.010 | 1.016 | 0.841 | 0.865 | 0.830 | 0.778 |
| Sveitsiske kvinner 2 | 0.665 | 0.691 | 0.677 | 0.689 | 0.315 | 0.330 | 0.330 | 0.277 |
| Islandske menn 2 | 2.097 | 1.326 | 1.319 | 1.327 | 2.215 | 1.585 | 1.567 | 1.591 |
| Islandske kvinner 2 | 1.834 | 1.854 | 1.865 | 1.869 | 1.888 | 1.943 | 1.930 | 1.918 |

* Post-glatting bare av b_x ettersom det viser seg å være vanskelig å finne passende startverdier til post-glatting av a_x .

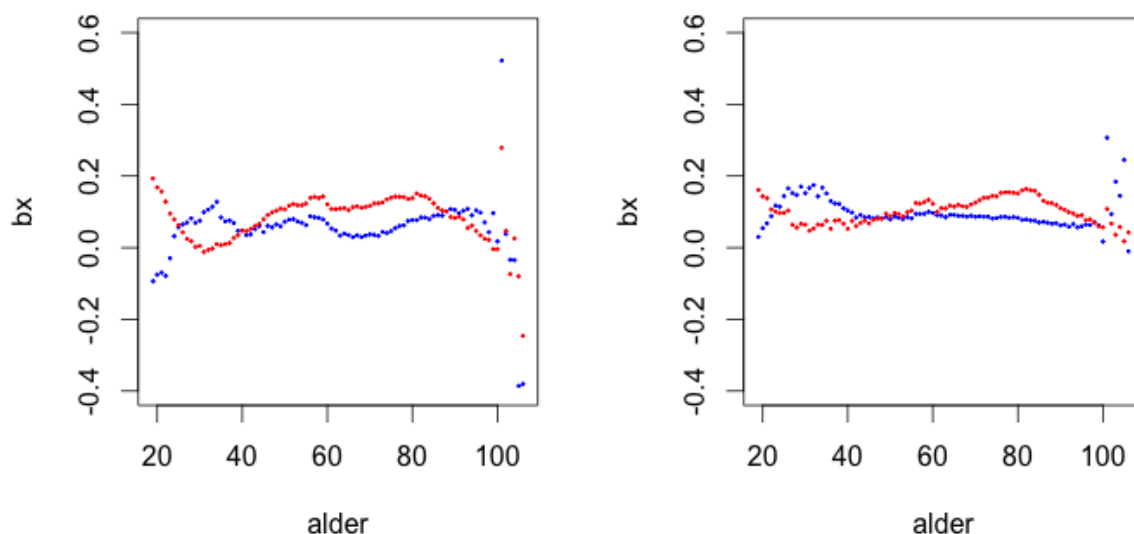
4.2 Et første inntrykk

At det er grunnleggende forskjeller mellom Lee-Carter modellen og det parametriske alternativet kommer tydelig til syne i fremskrivning av e_{18} . Av figur 21 (nederst til høyre) fremkommer det at både Lee-Carter (alle tre metodene) og den parametriske modellen underestimerer \hat{e}_{18k} i fremskrivningsperioden for sveitsiske menn i scenario 1. I samsvar med resultater i blant annet Booth et al. (2005) viser dette seg å være en generell trend for menn i scenario 1. Hovedgrunnen er antakelig at trenden i dødelighetsnedgangen for perioden 1950-1999 ikke er representativ (ikke bratt nok) for perioden 2000-2010. Det samme gjelder også scenario 2, men i en noe mindre grad for Island, se figur 24 og 26 for en illustrasjon av franske og islandske menn i scenario 2. Med unntak av sveitsiske og islandske menn i scenario 1 kommer den parametriske modellen best ut, se tabell 11 hvor $RMSE_{\hat{e}_{18}}^{frem}$ i dette tilfelle hovedsakelig gjenspeiler underestimering. Det samme gjelder for islandske kvinner i begge scenarier, den parametriske modellen er overlegen i fremskrivning av e_{18} og $RMSE_{\hat{e}_{18}}^{frem}$ gjenspeiler i stor grad underestimering for alle de tre versjonene av Lee-Carter modellen, se figur 25 (nederst til høyre) for en illustrasjon av scenario 2.

Tabell 11: $RMSE_{\hat{e}_{18}}$ fordelt på tilpasnings- og fremskrivningsperiode pr scenario, land og kjønn, under pre-, post- og uglattet Lee-Carter og den parametriske modellen, (Franske menn 2 er franske menn i scenario 2).

| | Tilpasning | | | | Fremskrivning | | | |
|----------------------|------------|--------|-------|-------|---------------|--------|-------|-------|
| | raw | post | pre | par | raw | post | pre | par |
| Franske menn 1 | 0.116 | 0.147 | 0.104 | 0.185 | 1.133 | 1.233 | 1.170 | 0.843 |
| Franske kvinner 1 | 0.155 | 0.167 | 0.179 | 0.136 | 0.186 | 0.194 | 0.306 | 0.195 |
| Sveitsiske menn 1 | 0.135 | 0.135* | 0.157 | 0.134 | 1.252 | 1.226* | 1.013 | 1.205 |
| Sveitsiske kvinner 1 | 0.158 | 0.193 | 0.262 | 0.144 | 0.117 | 0.130 | 0.288 | 0.120 |
| Islandske menn 1 | 0.141 | 0.217 | 0.405 | 0.398 | 1.400 | 1.583 | 0.588 | 1.115 |
| Islandske kvinner 1 | 0.215 | 0.294 | 0.348 | 0.360 | 1.319 | 1.444 | 1.137 | 0.766 |
| Franske menn 2 | 0.049 | 0.074 | 0.134 | 0.098 | 0.743 | 0.738 | 0.877 | 0.624 |
| Franske kvinner 2 | 0.056 | 0.162 | 0.139 | 0.104 | 0.513 | 0.408 | 0.488 | 0.721 |
| Sveitsiske menn 2 | 0.059 | 0.121 | 0.581 | 0.063 | 1.167 | 1.332 | 2.372 | 1.040 |
| Sveitsiske kvinner 2 | 0.057 | 0.136 | 0.063 | 0.126 | 0.751 | 0.501 | 0.500 | 1.193 |
| Islandske menn 2 | 0.281 | 0.448 | 0.398 | 0.374 | 0.674 | 0.561 | 0.582 | 0.481 |
| Islandske kvinner 2 | 0.282 | 0.361 | 0.318 | 0.291 | 0.605 | 0.565 | 0.423 | 0.363 |

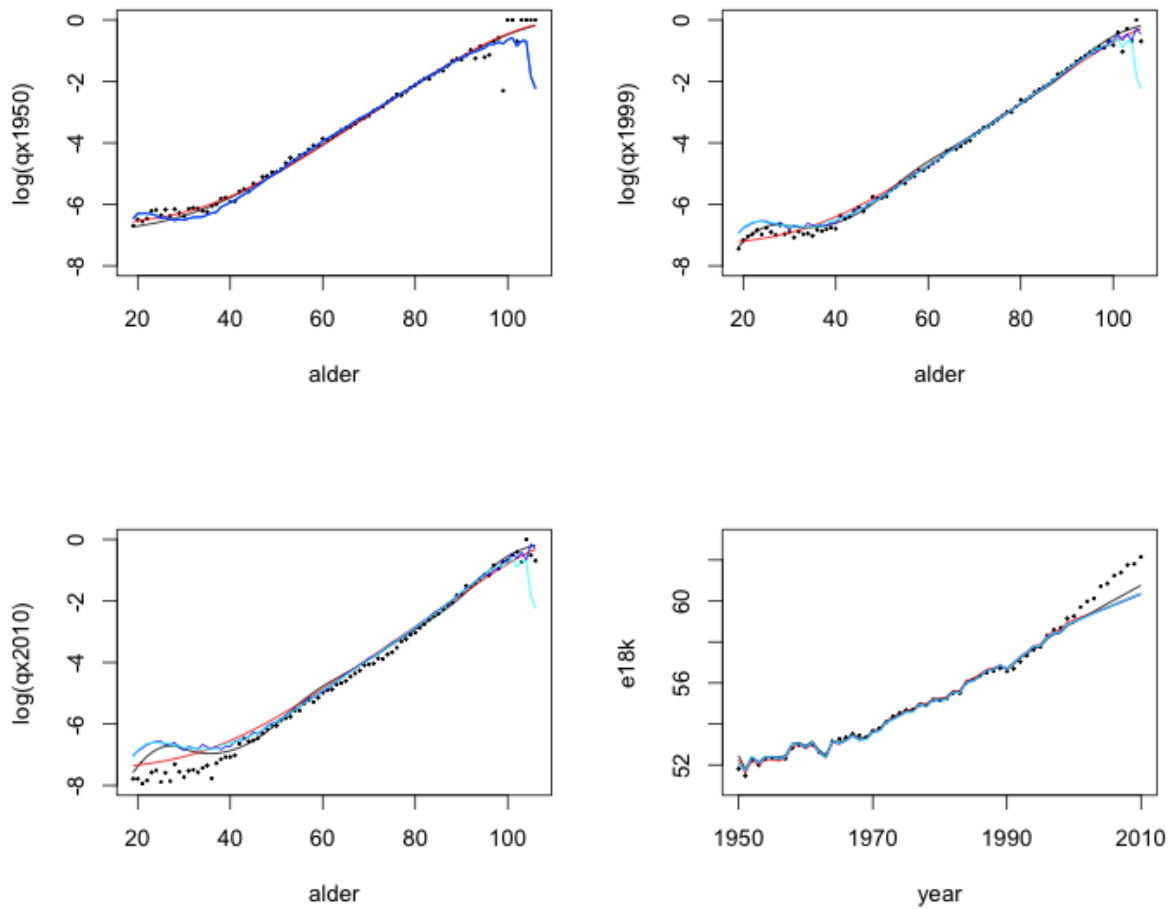
* Post-glatting bare av b_x ettersom det viser seg å være vanskelig å finne passende startverdier til post-glatting av a_x .



Figur 20: b_x , uglattet Lee-Carter, tilpasningsperiode 1950-1974 (i blått) og 1975-1999 (i rødt), franske menn (til venstre), franske kvinner (til høyre).

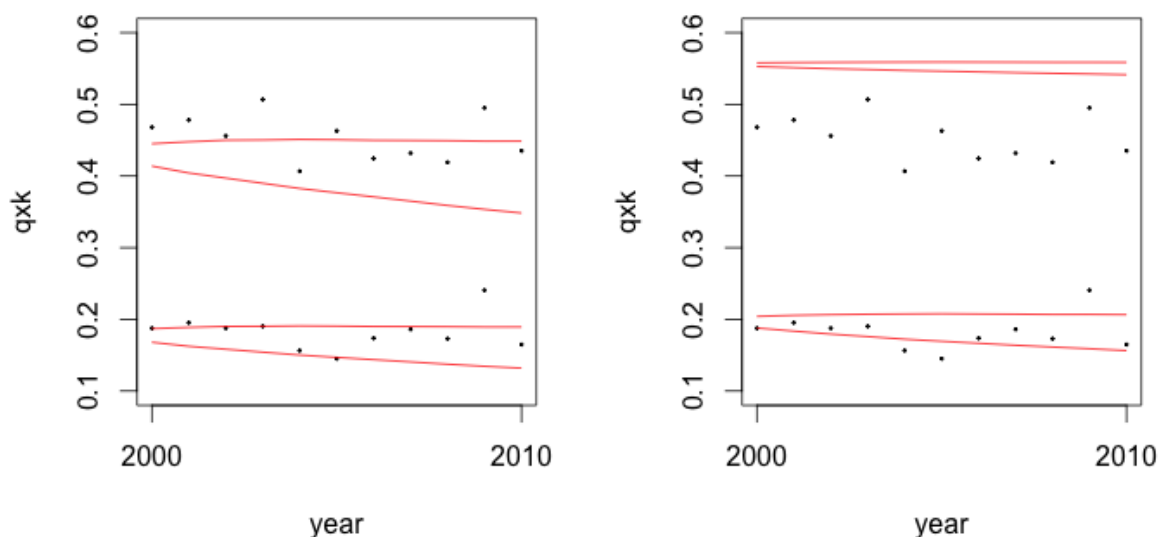
For franske og sveitsiske kvinner er situasjonen en annen. I scenario 1 er det lite skjevhet, trenden i perioden 1950-1999 ser ut til å være representativ for 2000-2010. Videre indikerer tabell 11 at valget mellom post- og uglattet Lee-Carter og den parametriske modellen er tilnærmet likegyldig, mens pre-glattet Lee-Carter presterer noe dårligere. I scenario 2 overestimeres \dot{e}_{18} i fremskrivningsperioden uansett modellvalg, men mest utpreget er det ved bruk av den parametriske modellen, se tabell 11 der $RMSE_{\dot{e}_{18}}^{frem}$ nå hovedsakelig representerer overestimering. Igjen skyldes dette sannsynligvis i stor grad at trenden i dødelighetsnedgangen for perioden 1970-1989 ikke er representativ (for bratt) for perioden 1990-2010.

Det er som tidligere nevnt kjent at Lee-Carter modellen har en tendens til å underestimere forventet gjenværende levetid. Lee and Miller (2001) legger frem en teori om at dette tildels kan skyldes at antakelsen om konstant b_x ikke holder, men gir ingen videre forklaring. Lee and Miller (2001) analyserer dødelighetsnedgangen for kvinner og menn samlet i en rekke land og viser at antakelsen om konstant b_x ikke holder for perioden 1900-1995, men antyder derimot at den er gjeldende for perioden 1950-1995. Figur 20 gir imidlertid en indikasjon på at dette ikke gjelder for franske menn og kvinner hver for seg, og tilsvarende figurer for Sveits og Island gir liknende bilder (ikke vist her). Absoluttverdien av b_x og ζ må sees i sammenheng, men ettersom ζ i alle tilfeller er positiv og større enn 0.1 antyder figur 20 at det kjønns spesifikke aldersmønsteret i dødelighetsnedgangen for Frankrike har undergått i hvert fall ett skifte i løpet av de siste 50 årene. Dette skiftet er mer alvorlig for menn, ettersom b_x har gått fra negativ til positiv for aldersgruppen 18-22, og motsatt for aldersgruppen 30-32, mens b_x holder seg positiv for kvinner. Brudd på antakelsen om konstant b_x kan dermed kanskje være med på å forklare hvorfor Lee-Carter



Figur 21: $\log(\hat{q}_{xk})$ sveitsiske menn (i svart stiplet), tilpasning eller fremskrivning i scenario 1 under parametrisk modell (i rødt), pre- (i svart), post- (i turkist) og uglattet (i blått) Lee-Carter, første baseår (øverst til venstre), siste baseår (øverst til høyre) og siste år av fremskrivningsperioden (nederst til venstre). e_{18k} (nederst til høyre).

sammenliknet med det parametriske alternativet i mange tilfeller underestimerer \dot{e}_{18} . I motsetning til Lee-Carter sikrer den parametriske modellen, uavhengig av datasettet, at q_{xk} er monoton funksjon av x , og unngår dermed de ofte observerte utfordringene assosiert med fleksibel aldersspesifikk dødelighetsutvikling. Datasettet bestående av franske menn eller kvinner i perioden 1950-1999 er relativt stort, selv innen demografi, og urealistisk stort i livsforsikringssammenheng. At det, for et datagrunnlag av denne størrelsen, er såpass stor usikkerhet rundt Lee-Carter modellens underliggende antakelser, kan tale for bruk av en enklere og mindre fleksibel modell, i hvert fall når datagrunnlaget begrenses.



Figur 22: *Franske menn, \hat{q}_{90k} og \hat{q}_{100k} , med fremskrivningsintervall (i rødt) basert på tilpasning til franske menn 18-105 år i perioden 1950-99, under den parametriske modellen (til venstre) og pre-glattet Lee-Carter (til høyre).*

4.3 Noen utvalgte eksempler

At en metode er overlegen en annen i tilpasning eller fremskrivning av de aldersspesifikke dødelighetene, trenger ikke bety at det samme gjelder e_{18k} , ettersom tilpasnings- og fremskrivningsfeil i aldersspesifikke dødeligheter potensielt kan kanselleres mot hverandre i beregningen av e_{18k} . At en metode er overlegen en annen i tilpasning, er heller ikke entydig med at det samme gjelder for fremskrivning. Tabell 10 og 11 gir flere eksempler, men bare i form av to aggregerte mål som i seg selv bare kan brukes som en indikasjon på modellenes prestasjon. Vi har derfor tatt for oss noen av datasettene, som også representerer andre, i noe mer detalj. Her fokuserer vi på overensstemmelse mellom punktestimater og virkelighet, og om modellens langsiktige dynamikk ser ut til å være biologisk rimelig.

Metodenes fremskrivningsevne i form av fremskrivningsfordeling er minst like viktig, men vil ikke, grunnet oppgavens omfangsbegrensning, utforskes i detalj her. En svakhet med Lee-Carter modellen må imidlertid nevnes. Under Lee-Carter modellen er bredden på fremskrivningsintervallene proporsjonal med b_x . Dødelighetsnedgangen har historisk sett vært lavest for de aller eldste. Dermed er også b_x som regel lavest for denne aldersgruppen, og glattingsrutinen i avsnitt 2.2 forsterker effekten (tvinger b_x mot null for de eldste uavhengig av opprinnelig verdi). Resultatet er smalere fremskrivningsintervall for de eldste, og dette strider med at det historisk sett observeres større usikkerhet i dødelighet for denne aldersgruppen, se blant annet Cairns et al. (2011a). Den parametriske modellen har ikke denne svakheten. Figur 22 viser et eksempel basert på franske menn, og illustrerer

hvordan det parametriske alternativet i motsetning til Lee-Carter modellen fanger opp at usikkerheten i \dot{q}_{100k} er større, og ikke mindre, enn usikkerheten i \dot{q}_{90k} .

4.3.1 Sveitsiske menn, scenario 1

Et datagrunnlag på 50 år med rundt 2 millioner menn pr år må betraktes som relativt stort i de fleste sammenhenger. Tabell 10 og 11 indikerer at det ikke er store forskjeller i modellenes tilpasningsevne, mens det ser ut til at det parametriske alternativet og den pre-glattede versjonen av Lee-Carter modellen gjør det litt bedre i fremskrivning av henholdsvis $\log(q_{xk})$ og e_{18k} . Figur 21 gir et samsvarende bilde.

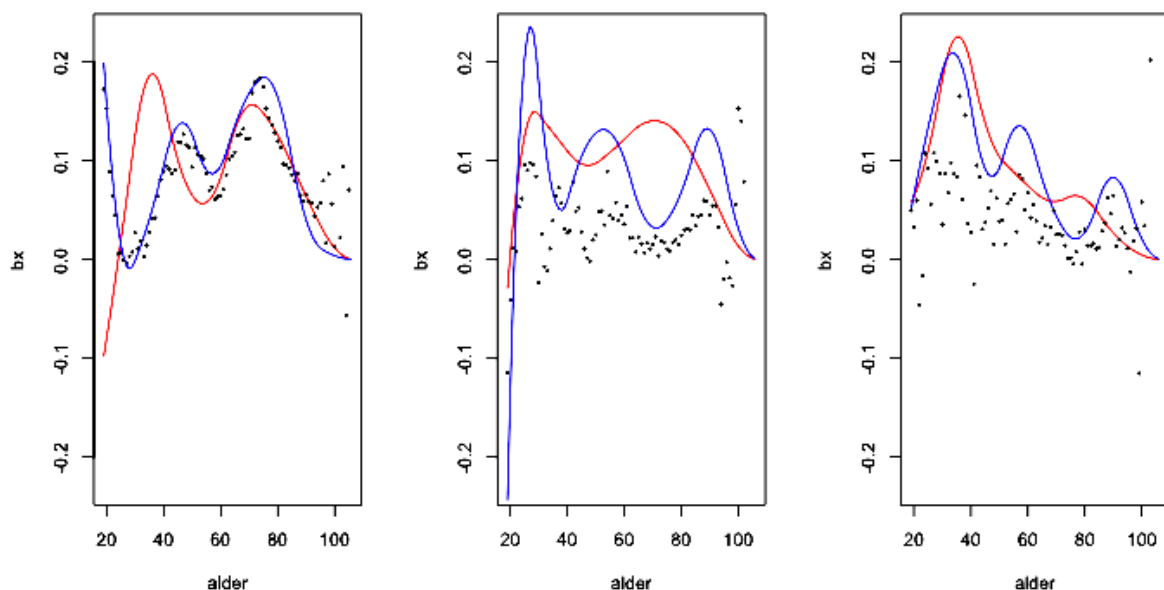
I første tilpasningsår kan effekten av glatting av a_x observeres, øverste venstre del av figur 21 viser hvordan pre-glatting jevner ut uregelmessighetene som er tilstede ved bruk av den uglattede versjonen av Lee-Carter modellen. Effekten av glatting av b_x er kanskje best synlig i siste fremskrivningsår, se nederst til venstre i figur 21. Figur 21 gir også et eksempel på en mulig konsekvens av post-glatting av kun b_x . Uten glatting av a_x er dødeligheten i første baseår identisk ved post- og uglattet Lee-Carter. For de eldste er uglattet b_x stor og negativ (se figur 19, men merk at verdier for de aller eldste er utelatt av hensyn til format), og underestimeringen av $\log(q_{x1950})$ korrigeres i løpet av tilpasningsperioden. Glatting tvinger b_x mot null for de eldste, se venstre side av figur 19, og resulterer i at underestimeringen av $\log(q_{x1950})$ forblir tilnærmet uendret ut tilpasnings- og fremskrivningsperioden.

Forskjeller mellom det mindre fleksible parametriske alternativet og Lee-Carter modellen er best synlig i alderen 18-35, ettersom dødelighetskurven skifter form i løpet av tilpasningsperioden for denne aldersgruppen. Den mer fleksible Lee-Carter modellen fanger opp dette skiftet, og spesielt den pre-glattede versjonen gir en bedre tilpasning i siste baseår, se øverst til høyre i figur 21. Den observerte formforandringen viser seg imidlertid å være flyktig, og i siste fremskrivningsår er det den mindre fleksible parametriske modellen som gir det mest realistiske bilde av dødeligheten for den aktuelle aldersgruppen.

4.3.2 Franske menn, scenario 2

En baseperiode på 20 år kan ikke klassifiseres som lang i demografisk sammenheng, Lee and Carter (1992) observerte ustabiliteter i fremskrivningene da deres opprinnelige modell ble tilpasset den amerikanske populasjonen med en baseperiode på 10-20 år. I livsforsikringssammenheng er derimot 20 år med data relativt mye, og 18 millioner er et urealistisk stort antall pr år. Tabell 10 og 11 antyder at den uglattede Lee-Carter modellen gir best tilpasning mens det parametriske alternativet er bedre på fremskrivning. Figur 24 gir et tilsvarende bilde.

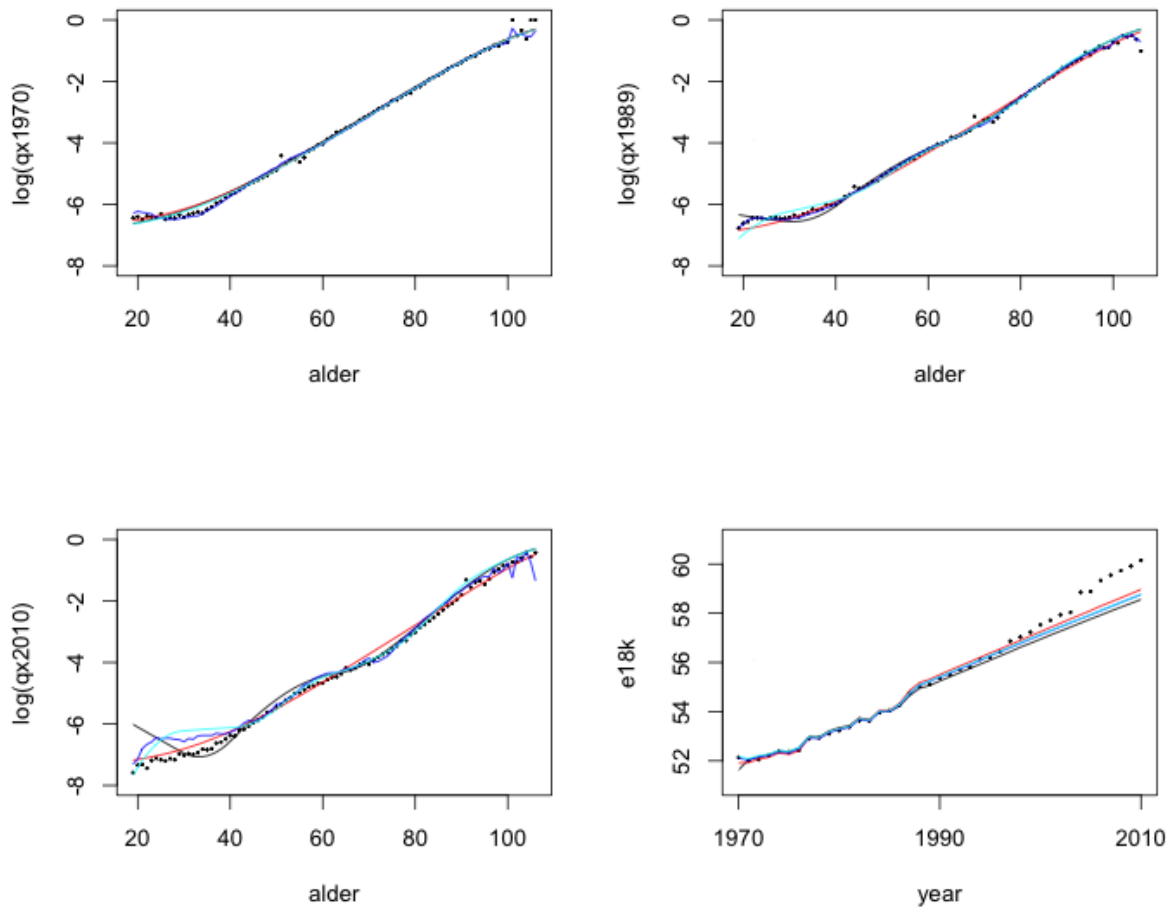
Som i eksempelet med sveitsiske menn i scenario 1, ser det ut til at dødelighetskurven har gjennomgått et formskifte i løpet av tilpasningsperioden, denne gangen for aldersgruppen 18-30. I dette tilfelle ser det imidlertid ut som det bare er den uglattede versjonen som fanger opp hele skiftet, mens de post- og pre-glattede versjonene kun ser ut til å overdrive formskiftet for henholdsvis første og siste halvpart av den aktuelle aldersgruppen. Igjen



Figur 23: b_x , pre- (i rødt), post- (i blått) og uglattet (i svart) Lee-Carter, tilpasningsperiode 1970-89, franske menn (til venstre), islandske kvinner (i midten) og islandske menn (til høyre).

ser skiftet ut til å være flyktig og den parametriske modellen gir igjen, for denne aldersgruppen, et mer realistisk bilde av virkelig dødelighet i siste fremskrivningsår.

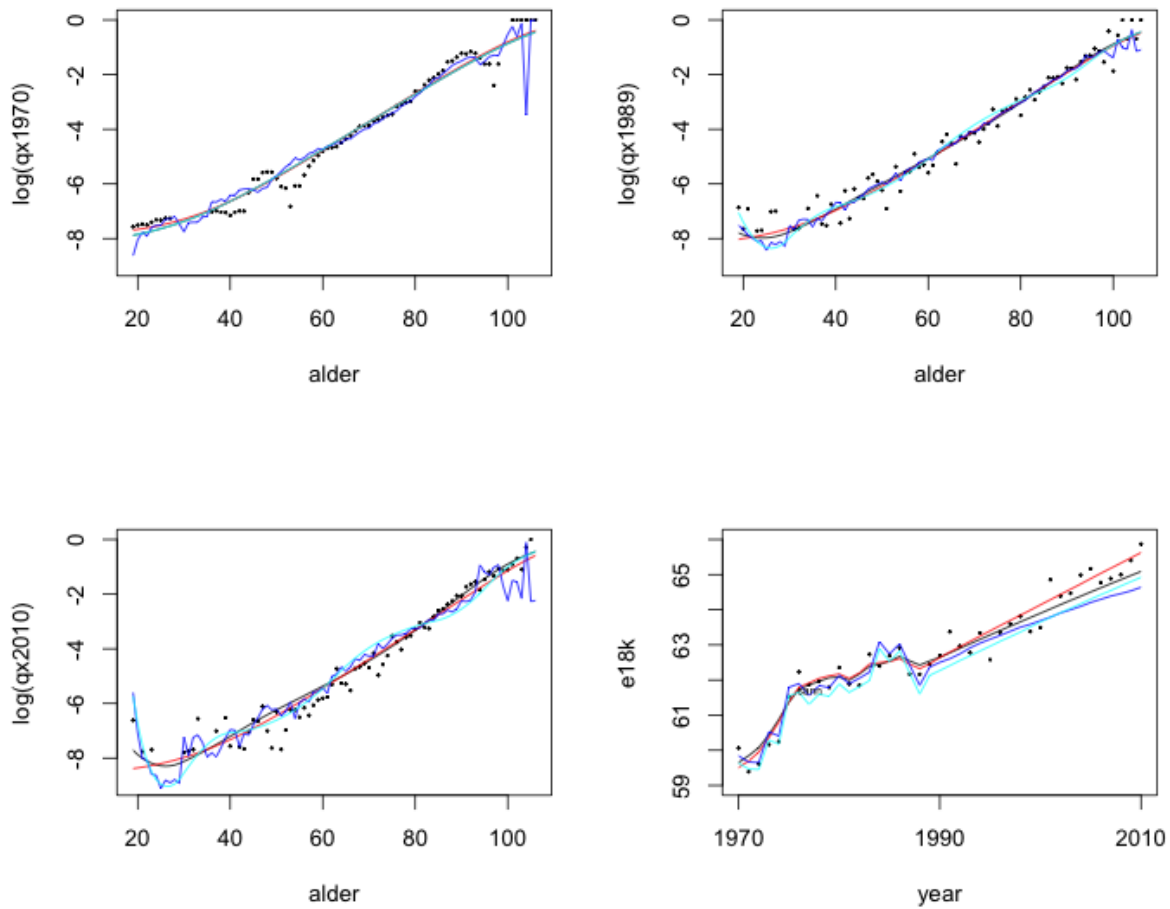
Ettersom ζ og σ i dette tilfellet er tilnærmet like under de tre versjonene av Lee-Carter modellen, finnes forklaringen på de tre versjonenes ulike oppførselen ved å kombinere tilpasningen i første baseår, se figur 24, med formen på b_x -kurvene i venstre del av figur 23. Både den pre- og post-glattede versjonen overestimerer dødeligheten i første baseår for aldersgruppen 25-35. Den uglattede versjonen fanger opp at den faktiske dødeligheten stiger, negativ b_x , for 25- og 26-åringer i løpet av tilpasningsperioden, men dette viser seg altså å være en flyktig trend, se siste fremskrivningsår i figur 24. Både post- og pre-glatting resulterer i at denne flyktige trenden spres, og med pre-glatting også forflyttes, til aldersgrupper som i virkeligheten har opplevd en dødelighetsnedgang i løpet av tilpasningsperioden. Post-glatting gir negativ b_x for aldersgruppen 26-29, mens pre-glatting resulterer i negativ b_x for aldersgruppen 18-23, se venstre side av figur 23. Figur 19 viser to liknende tilfeller med pre-glatting, omtrentlig samme form på den uglattede b_x -kurven, og pre-glatting er utfordrende (venstre side) eller resulterer som her i urimelige verdier for de yngste (høyre side). Årsaken er antakeligvis flate likelihoodoverflater, se diskusjon i tilknytning til figur 19.



Figur 24: $\log(\dot{q}_{xk})$ franske menn (i svart stiplet), tilpasning eller fremskrivning i scenario 2 under parametrisk modell (i rødt), pre- (i svart), post- (i turkist) og uglattet (i blått) Lee-Carter, første baseår (øverst til venstre), siste baseår (øverst til høyre) og siste år av fremskrivningsperioden (nederst til venstre). e_{18k} (nederst til høyre).

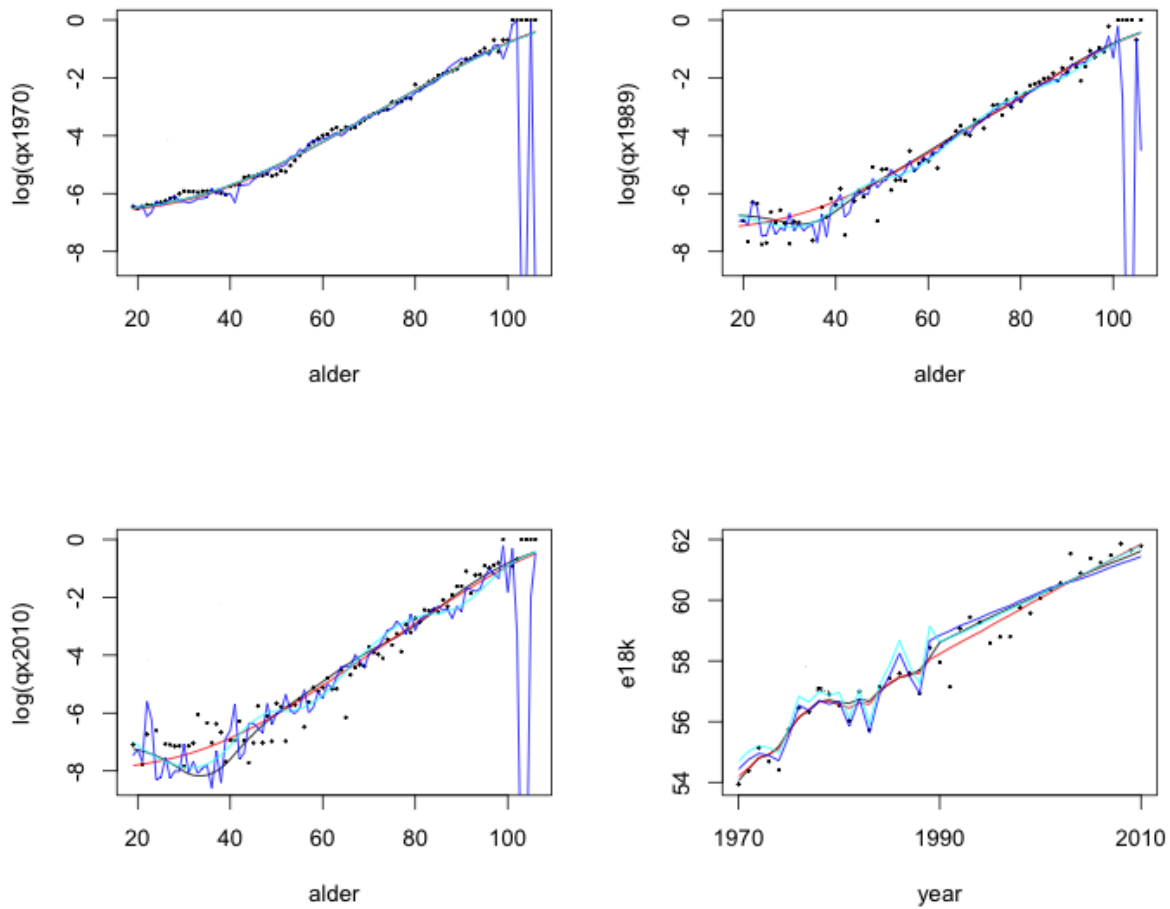
4.3.3 Islandske kvinner og menn, scenario 2

Et datagrunnlag på 20 år med rundt 70 000 personer pr år vil anses som begrenset innen demografi, men er nok mer representativt for en livsforsikringsportefølje enn de andre datasettene vi har tilgjengelig, og vi velger derfor å se nærmere på både islandske kvinner og menn i scenario 2. Tabell 10 indikerer at det ikke er store forskjeller mellom metodene i tilpasning og fremskrivning av $\log(q_{xk})$. For menn ser det imidlertid ut til at den uglattede versjonen av Lee-Carter modellen presterer noe dårligere, mens det motsatte er tilfellet for kvinner. Videre antyder tabell 11 at den uglattede versjonen av Lee-Carter modellen tilbyr en noe bedre tilpasning av e_{18k} , mens det parametriske alternativet er best for fremskrivning.



Figur 25: $\log(\hat{q}_{xk})$ islandske kvinner (i svart stiptet), tilpasning eller fremskrivning i scenario 2 under parametrisk modell (i rødt), pre- (i svart), post- (i turkist) og uglattet (i blått) Lee-Carter, første baseår (øverst til venstre), siste baseår (øverst til høyre) og siste år av fremskrivningsperioden (nederst til venstre). e_{18k} (nederst til høyre).

Figur 25 og 26 gir et litt annet bilde. I samsvar med tabell 11 ser den parametriske modellen ut til å gi et riktigere bilde av den langsiktige dynamikken (se utvikling i e_{18} nederst til høyre), men med et datagrunnlag av denne størrelsen er glatting, i motsetning til hva tabell 10 og 11 kan antyde, av a_x og spesielt b_x nødvendig. Det kan sjeldent være ønskelig å modellere dødeligheten som en så volatil og uregelmessig prosess som figur 25 og 26 viser at en uglattet Lee-Carter gjør i disse tilfellene. Figur 23 illustrerer de underliggende uregelmessighetene i b_x . Merk at graden av glatting nå er betydelig større ved ett-trinns estimering, og legg også merke til hvordan svingningene i den post-glattede b_x -kurven forplanter seg til tilpasset og fremskrevet dødelighet og forventet gjenværende levetid ved alder 18 i figur 25 og 26.



Figur 26: $\log(\hat{q}_{xk})$ islandske menn (i svart stiplet), tilpasning eller fremskrivning i scenario 2 under parametrisk modell (i rødt), pre- (i svart), post- (i turkist) og uglattet (i blått) Lee-Carter, første baseår (øverst til venstre), siste baseår (øverst til høyre) og siste år av fremskrivningsperioden (nederst til venstre). e_{18k} (nederst til høyre).

Det viser seg, kanskje ikke så overraskende, at graden av glatting ved ett- og to-trinns estimering avhenger av størrelsen på datagrunnlaget. Figur 19 og 23 gir en illustrasjon, med et stort datagrunnlag, og dermed forholdsvis regelmessige uglattede b_x -estimer, antyder figur 19 at graden av glatting er høyere med to-trinns estimering, mens figur 23 viser hvordan ett-trinns estimering gir større glattingseffekt ettersom datagrunnlaget begrenses. I midtre og høyre del av figur 23 kan det også se ut som det har oppstått en skaleringsfeil, de pre-glattede kurvene ligger konsekvent over de uglattede estimatene, men tabell 12 gir forklaringen. Dette likner på situasjonen vist i figur 3 og 8 og tabell 2 og 6 i avsnitt 3.1. I avsnitt 3.1 viste det seg at etter som datagrunnlaget begrenses, har både post- og særlig uglattet Lee-Carter en tendens til å overestimere usikkerheten i fremtidig dødelighet, og tabell 12 kan kanskje antyde at dette er tilfelle også for virkelige data.

Som i de to foregående eksemplene observeres det et skifte av dødelighetskurvens form i

løpet av tilpasningsperioden for menn i aldersgruppen 18-40, se figur 26. Både post- og pre-glatting forhindrer i dette tilfelle, i motsetning til den uglattede versjonen, at aldersspesifikke trender med økning i dødeligheten i løpet av tilpasningsperioden fanges opp, se høyre side av figur 23. Det ser likevel ut til at både post- og pre-glatting fanger opp hovedlinjene i formskiftet og gir dermed, for denne aldersgruppen, en bedre tilpasning enn den parametriske modellen i siste baseår. Siste fremskrivningsår i figur 26 indikerer, på samme måte som i foregående eksempler, at formskiftet var flyktig, og den mindre fleksible parametriske modellen gir igjen et mer realistisk bilde av dødelighetsutviklingen for denne aldersgruppen. Figur 26 antyder imidlertid også at glatting i dette tilfelle forhindret at tilfeldige trender ble fanget opp. I fremskrivningsperioden er det andre aldre enn i tilpasningsperioden som opplever økt dødelighet, legg merke til den heller dårlige treffprosenten til den uglattede versjonen, med sine spisse topper og bunner, i siste fremskrivningsår.

For islandske kvinner forhindrer ikke glatting modellering av $b_x < 0$, men pre-glatting begrenser utbredelsen til 18-åringer og post-glatting til aldersgruppen 18-20. Ettersom verdien på ζ er betydelig lavere ved post- og pre-glatting, vil den langsiktige effekten, spesielt med pre-glatting, være dempet sammenliknet med den uglattede versjonen, se siste fremskrivningsår i figur 25 for en indikasjon. I dette tilfelle ser det ut til at det er en svak økning i faktisk dødelighet også gjennom fremskrivningsperioden for 18- og 20-åringer, men ikke for 19-åringer, noe som kan indikere at det dreier seg om tilfeldige svingninger. Det parametriske alternativet gir nok en gang det mest realistiske bildet av den langsiktige dødelighetsutviklingen for denne aldersgruppen.

Tabell 12: *Dynamiske parametre, ζ og σ , islandske kvinner og menn i scenario 2, pre-, post- og uglattet (raw) Lee-Carter.*

| | Kvinner | | | Menn | | |
|----------|---------|-------|-------|-------|-------|-------|
| | pre | post | raw | pre | post | raw |
| ζ | 0.125 | 0.199 | 0.563 | 0.225 | 0.261 | 0.592 |
| σ | 0.255 | 0.705 | 1.993 | 0.333 | 0.915 | 2.071 |

4.4 Oppsummering

Mangel på glatting av Lee-Carters modell fører til uregelmessigheter i tilpasset og fremskrevet $\log(q_{xk})$ selv med 50 år med data og rundt 2 millioner menn pr år, sveitsiske menn i scenario 2. Glattingseffekten øker med minkende datagrunnlag, og for islandske kvinner og menn er den uglattede versjonen så volatil at den anses som biologisk urimelig. Det observeres imidlertid utfordringer i forbindelse med glatting av b_x for aldersgruppen 18-40. Både post- og pre-glatting viser seg å kunne resultere i at en trend med økning i faktisk dødelighet i tilpasningsperioden spres eller forflyttes til aldersgrupper som i virkeligheten har opplevd en dødelighetsnedgang. Dette fører til biologisk urimelig langsiktig dynamikk for den berørte aldersgruppen. En mulighet kan være å sette begrensningen $b_x \geq 0$ for datasett av mindre størrelse, men da forsvinner imidlertid noe av poenget med fleksibel, aldersspesifikk dødelighetsutvikling.

Selv om glatting begrenser Lee-Carter modellens fleksibilitet, er spline-funksjonen fortsatt fleksibel nok til å tillate modellering av aldersspesifikke skift av dødelighetskurvens form i løpet av tilpasningsperioden. Resultatet er som oftest en bedre tilpasning i siste baseår enn den mindre fleksible parametriske modellen kan tilby. Det viser seg imidlertid at slike trender, også for større datasett, kan være flyktige (b_x i tilpasningsperioden representerer ikke b_x i fremskrivningsperioden, eller antakelsen om konstant b_x i tilpasningsperioden holder ikke), og at Lee-Carter modellens fleksibilitet dermed kan resultere i en heller urealistisk modellering av langsiktig dynamikk for den berørte aldersgruppen. Konsekvensen er naturlig nok større for mindre populasjoner, som ofte preges av tilfeldige variasjoner. Det mindre fleksible parametriske alternativet forhindrer, gjennom modellering av q_{xk} som en monoton funksjon av x , at aldersspesifikke tilfeldige eller flyktige trender fanges opp, og den langsiktige dynamikken under modellen ser i de fleste tilfeller, og spesielt for mindre populasjoner, ut til å være mer realistisk enn under noen av de tre versjonene av Lee-Carter modellen. Lee-Carter type modellering av aldersspesifikk dødelighet fører også til biologisk urimelige fremskrivningsintervall, en svakhet som ikke deles med den parametriske modellen.

Effekten av ett-trinns estimering kommer først til syne når datagrunnlaget er såpass begrenset som det er for islandske kvinner og menn. Virkningen av glatting av b_x er betydelig større enn ved to-trinns estimering, og forplantes til tilpasset og fremskrevet $\log(q_{xk})$ og e_{18k} . Sett i sammenheng med avsnitt 3.1 ser det ut til at ett-trinns estimering (pre-glatting) av Lee-Carter modellens parametre gir et mer realistisk bilde av usikkerhet i fremtidig dødelighet. Det viser seg imidlertid at pre-glatting av Lee-Carter modellen kan være assosiert med spesielt flate likelihoodoverflater. Dette kan gi opphav til situasjoner, som den for franske menn i scenario 1, der den pre-glattede modellen forespeiler en økning i dødeligheten for de yngste selv om trenden i datasettet tilsier en nedgang.

5 Konklusjon og videre arbeid

I denne oppgaven har sammenhengen mellom modellens parametriseringsgrad og modellens evne til å håndtere tilfeldig variasjon langt på vei blitt etablert. Et mindre fleksibelt parametrisk alternativ til Lee-Carter type modeller har blitt presentert og videreutviklet. Gjennom simuleringseksperimenter antydes det at denne modellen er overlegen en pre-glattet Lee-Carter i håndtering av tilfeldig variasjon, og den resulterende reduksjonen i skjevhet og usikkerhet for engangspremier ser ut til å kunne være betydelig for beregning av fremtidige forpliktelser i en enkelt pensjonskasse, eller for prising av livsvarige annuiteter.

Gjennom en ex-post analyse av tilpasnings- og fremskrivningsevne antydes det at det parametriske alternativet i mange sammenhenger kan gi et mer realistisk bilde av fremtidig dødelighet enn Lee-Carter type modeller. Dette gjelder spesielt for små populasjoner, som ofte preges av tilfeldige variasjoner, og grunner sannsynligvis i at Lee-Carter modellens fleksibilitet muliggjør modellering av aldersspesifikke trender som kan være flyktige eller, for små populasjoner, tilfeldige. Lee-Carter type modellering av aldersspesifikk dødelighet fører også til biologisk urimelige fremskrivningsintervall. Den parametriske modellen har ikke denne svakheten. Det er imidlertid ikke foretatt noen analyse av fremskrivningsfordeling i denne oppgaven, noe som kunne vært interessant å studere nærmere ved en annen anledning. Før den parametriske modellen eventuelt tas i bruk, bør den studeres nærmere, for å blant annet fastsette optimale parameterbegrensninger.

Ett-trinns estimeringsmetoden presentert i denne oppgaven er, som andre ett-trinns metoder, relativt implementeringskrevende sammenliknet med tradisjonell to-trinns estimering. Gjennom simuleringseksperimenter antydes det imidlertid at i situasjoner der et relevant datagrunnlag er begrenset, kan fordelene mer enn oppveie for implementeringskravet. Pre-glattning av Lee-Carters modell ser ut til å gi et mer realistisk bilde av usikkerhet i fremtidig dødelighet, og virkningen ser ut til å øke etter som datagrunnlaget begrenses. Videre antydes det at hvis Lee-Carter type modeller skal anvendes for beregning av fremtidige forpliktelser i en enkelt pensjonskasse, eller for prising av livsvarige annuiteter, kan ett-trinns estimeringsmetoden redusere skjevhet og risiko i en såpass stor grad at det er snakk om økonomiske konsekvenser.

Det viser seg imidlertid at pre-glattning av Lee-Carters modell kan være assosiert med spesielt flate likelihoodoverflater. Dette kan gi opphav til situasjoner der den pre-glattede modellen forespeiler en økning i dødeligheten for de yngste selv om trenden i datasettet tilsier en nedgang for alle aldre. En mulig løsning kan kanskje være å øke tettheten av spline-funksjonens vendepunkter rundt de yngre aldersgruppene. Hvis dette ikke er tilstrekkelig, er kanskje løsningen en mer komplisert glattingsrutine som muliggjør for eksempel valg av optimale vendepunkter og enderestriksjoner. Da ender vi imidlertid kanskje opp, som i Zhao (2012), med en utfordrende oppgave innen modellvalg, og metodens brukervennlighet vil kanskje reduseres. Dette er ikke utforsket i denne oppgaven, men det anbefales at problemet studeres nærmere før metoden eventuelt tas i bruk.

Videre fremkommer det at både post- og pre-glattning kan resultere i at en trend med

økning i faktisk dødelighet i tilpasningsperioden spres eller forflyttes til aldersgrupper som i virkeligheten har opplevd en dødelighetsnedgang. Dette kan gi opphav til biologisk urimelig langsiktig dynamikk for den berørte aldersgruppen. En mulighet kan være å sette begrensningen $b_x \geq 0$ for datasett av mindre størrelse, men da forsvinner imidlertid noe av poenget med fleksibel, aldersspesifikk dødelighetsutvikling. Kanskje er det på tide å la Lee-Carter modellen være det den ble utviklet til, en modell først og fremst for den amerikanske populasjonen og populasjoner med liknende størrelse og egenskaper, og heller la seg inspirere til å utvikle og ta i bruk modeller som er tilment et begrenset datagrunnlag.

Helt tilslutt vender vi tilbake til et av spørsmålene som ble tatt opp innledningsvis: Hva er en rettfærdig pris på en livsvarig annuitet? Denne oppgaven gir ikke noe konkret svar, men antyder at det kan stilles spørsmål ved at en del av vårt nåværende lovpålagte beregningsverktøy, K2013, er basert på en fleksibel og uglattet versjon av Lee-Carter modellen med lite forankrede sikkerhetsmarginer. For beregning av fremtidige forpliktelser, og spesielt prising av livsvarige annuiteter og MLS, er det kanskje på tide å rette fokus mot fremskrivningsevne heller enn tilpasningsevne eller modellstabilitet heller enn fleksibilitet. Hvis det private annuitetsmarkedet er så oppadgående som Lin and Cox (2005) spår, er det antakeligvis bare et spørsmål om tid før kundene krever mer individuell prising, og for et begrenset datagrunnlag indikerer denne oppgaven at mindre fleksible modeller, som det parametriske alternativet presentert her, kan ha potensial.

Referanser

- Andrieu, C., Doucet, A., and Punskeya, E. (2001). Sequential monte carlo methods for optimal filtering. In *Sequential Monte Carlo Methods in Practice*. Springer, New York, second edition.
- Arulampalam, M. S., Maskell, S., Gordon, N., and Clapp, T. (2002). A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. *IEEE Transactions on Signal Processing*, 50:174–188.
- Baltzersen, M. and Steffensen, E. (2013). Nytt dødelighetsgrunnlag i kollektiv pensjonsforsikring (k2013). Brev fra finanstilsynet til alle livsforsikringsselskap og pensjonskasser datert 8. mars 2013. Available at <http://www.finanstilsynet.no>.
- Bell, W. (1997). Comparing and assessing time series methods for forecasting age-specific fertility and mortality rates. *Journal of Official Statistics*, 13:279–303.
- Berntsen, K. (2013). Forventet levealder i oslo bydeler: Fortsatt store forskjeller i levealder i oslo. *Population Studies*, 4:18–26.
- Booth, H., Hyndman, R., Tickle, L., and de Jong, P. (2006). Lee-carter mortality forecasting: a multi-country comparison of variants and extensions. *Demographic Research*, 15:289–310.
- Booth, H., Maindonald, J., and Smith, L. (2002). Applying lee-carter under conditions of variable mortality decline. *Population Studies*, 56:325–336.
- Booth, H. and Tickle, L. (2008). Mortality modelling and forecasting: A review of methods. *Annals of Actuarial Science*, 3:3–43.
- Booth, H., Tickle, L., and Smith, L. (2005). Evaluation of the variants of the lee-carter method of forecasting mortality: A multi- country comparison. *New Zealand Population Review*, 31:13–34.
- Brouhns, N., Denuit, M., and Keilegom, I. V. (2005). Bootstrapping the poisson log-bilinear model for mortality forecasting. *Scandinavian Actuarial Journal*, 2005, 3:221–224.
- Brouhns, N., Denuit, M., and Vermunt, J. (2002a). Measuring the longevity risk in mortality projections. *Bulletin of the Swiss Association of Actuaries*, 2002, 2:105–130.
- Brouhns, N., Denuit, M., and Vermunt, J. (2002b). A poisson log-bilinear regression approach to the construction of projected life tables. *Insurance: Mathematics and Economics*, 31:373–393.
- Bølviken, E. (2013). Dynamic mortalities by optimal non-linear filtering. Submitted Annals of Applied Statistics.
- Bølviken, E. and Moe, A. (2008). Gablers nye beregningsgrunnlag gap07: Metode og kontantverdier. Lecture notes in STK4520.

- Bølviken, E. and Storvik, G. (2001). Deterministic and stochastic particle filters in state space models. In *Sequential Monte Carlo Methods in Practice*. Springer, New York, second edition.
- Cairns, A., Blake, D., and Dowd, K. (2006a). Pricing death: Frameworks for the valuation and secularization of mortality risk. *Astin Bulletin*, 36:79–120.
- Cairns, A., Blake, D., and Dowd, K. (2006b). A two-factor model for stochastic mortality with parameter uncertainty: Theory and calibration. *The Journal of Risk and Insurance*, 73:687–718.
- Cairns, A., Blake, D., and Dowd, K. (2008). Modelling and management of mortality risk: A review. *Heriot-Watt University, and Pensions Institute Discussion Paper Pl-0814*.
- Cairns, A., Blake, D., Dowd, K., Coughlan, G., Epstein, D., and Khalaf-Allah, M. (2011a). Mortality density forecasts: An analysis of six stochastic mortality models. *Insurance: Mathematics and Economics*, 48:355–367.
- Cairns, A., Blake, D., Dowd, K., Coughlan, G., Epstein, D., Ong, M. A., and Balevich, I. (2009). A quantitative comparison of stochastic mortality models using data from england and wales and the united states. *North American Actuarial Journal*, 13:1–35.
- Cairns, A., Blake, D., Dowd, K., Coughlan, G., and Khalaf-Allah, M. (2011b). Bayesian stochastic mortality modelling for two populations. *Astin Bulletin*, 41:29–59.
- Cossette, H., Delwarde, A., Denuit, M., Guillot, F., and Marceau, E. (2007). Pension plan valuation and mortality projection a case study with mortality data. *North American Actuarial Journal*, 11:1–34.
- Cressie, N. and Wikle, C. (2011). *Statistics for Spatio-Temporal Data*. Wiley, New Jersey.
- Currie, I. D. (2013). Smoothing constrained generalized linear models with an application to the lee-carter model. *Statistical modelling*, 13:69–93.
- Czado, C., Delwarde, A., and Denuit, M. (2005). Bayesian poisson log-bilinear mortality projections. *Insurance: Mathematics and Economics*, 36:260–284.
- Delwarde, A., Denuit, M., and Eilers, P. (2007a). Smoothing the lee-carter and poisson log-bilinear models for mortality forecasting. a penalized log-likelihood approach. *Statistical modelling*, 7:29–48.
- Delwarde, A., M.Denuit, and Patrat, C. (2007b). Negative binomial version of the lee-carter model for mortality forecasting. *Applied Stochastic Models in Business and Industry*, 23:385–401.
- Doucet, A., de Freitas, N., and Gordon, N. (2001). An introduction to sequential monte carlo methods. In *Sequential Monte Carlo Methods in Practice*. Springer, New York, second edition.
- Dowd, K., Blake, D., Cairns, A., and P.Dawson (2006). Survivor swaps. *Journal of Risk and Insurance*, 73:1–17.

- Dowd, K., Cairns, A., Blake, D., Coughlan, G., Epstein, D., and Khalaf-Allah, M. (2010). Backtesting stochastic mortality models: An ex post evaluation of multi-period-ahead density forecasts. *North American Actuarial Journal*, 14:281–298.
- Gerber, H. (1997). *Life Insurance Mathematics*. Springer-Verlag Berlin Heidelberg, Germany, third edition.
- Giroi, F. and King, G. (2008). *Demographic Forecasting*. Princeton University Press, New Jersey.
- HMD (2013). Human mortality database. University of California, Berkeley (USA), and Max Planck Institute for Demographic Research (Germany). Available at www.mortality.org. Data downloaded april 2013.
- Hollmann, F., Mulder, T., and Kallan, J. (2000). Methodology and assumptions for the population projections of the united states. Population Division Working Paper, No. 38. Available at <http://www.census.gov/population/www/documentation/twps0038.pdf>.
- Jarner, S. and Kryger, E. (2011). Modelling adult mortality in small populations: The saint model. *Astin Bulletin*, 41:377–418.
- Jong, P. D. and Tickle, L. (2006). Extending lee-carter mortality forecasting. *Mathematical Population Studies*, 13:1–18.
- Keilman, N. and Pham, D. (2005). Hvor lenge kommer vi til å leve? levealder og aldersmønster for dødeligheten i norge, 1900-2060. *Økonomiske analyser*, 6:43–49.
- Kitagawa, G. (1987). Non-gaussian state-space modeling of non-stationary time series. *Journal of the American Statistical Association*, 82:1032–1041.
- Koissi, M.-C., Shapiro, A., and Høgnås, G. (2006). Evaluating and extending the lee-carter model for mortality forecasting: Bootstrap confidence interval. *Insurance: Mathematics and Economics*, 38:1–20.
- Lane, M. (2000). Pricing risk transfer transactions. *Astin Bulletin*, 30:259–293.
- Lee, R. and Carter, L. (1992). Modeling and forecasting u.s. mortality. *Journal of the American Statistical Association*, 87:659–671.
- Lee, R. and Miller, T. (2001). Evaluating the performance of the lee-carter method for forecasting mortality. *Demography*, 38:537–549.
- Li, J., Hardy, M., and Tan, K. (2009). Uncertainty in model forecasting: An extension to the classic lee-carter approach. *ASTIN Bullitin*, 39:137–164.
- Lin, Y. and Cox, S. (2005). Securitization of mortality risks in life annuities. *Journal of risk and Insurance*, 72:227–252.
- Pedroza, C. (2006). A bayesian forecasting model: predicting u.s. male mortality. *Biostatistics*, 7:530–550.

- Pitacco, E., Denuit, M., Haberman, S., and Oliveri, A. (2009). *Modelling Longevity Dynamics for Pensions and Annuity Business*. Oxford University Press, Oxford.
- Press, W., Teukolsky, S., Vetterling, W., and Flannery, B. (2007). *Numerical Recipes*. Cambridge University Press, Cambridge, third edition.
- Renshaw, A. and Haberman, S. (2003a). Lee-carter mortality forecasting with age-specific enhancement. *Insurance: Mathematics and Economics*, 33:255–272.
- Renshaw, A. and Haberman, S. (2003b). On forecasting of mortality reduction factors. *Insurance: Mathematics and Economics*, 32:379–401.
- Renshaw, A. and Haberman, S. (2006). A cohort-based extension to the lee-carter model for mortality reduction factors. *Insurance: Mathematics and Economics*, 38:556–570.
- Renshaw, A. and Haberman, S. (2008a). On simulation-based approaches to risk measurement in mortality with specific reference to binomial lee-carter modeling. In Society of Actuaries Living to 100 Symposium.
- Renshaw, A. and Haberman, S. (2008b). On simulation-based approaches to risk measurement in mortality with specific reference to poisson lee-carter modeling. *Insurance: Mathematics and Economics*, 42:797–816.
- Richards, S. J. and Currie, I. D. (2009). Longevity risk and annuity pricing with the lee-carter model. *British Actuarial Journal*, 15:317–343.
- Ross, S. (2010). *Introduction to Probability Models*. Academic Press, Burlington, tenth edition.
- Tuljapurkar, S., Li, N., and Boe, C. (2000). A universal pattern of mortality decline in the g7 countries. *Nature*, 405:789–792.
- Wei, W. (2006). *Time Series Analysis Univariate and Multivariate Methods*. Pearson Education, USA, second edition.
- Wilmoth, J. (1993). Computational methods for fitting and extrapolating the lee-carter model of mortality change. Technical report, Department of Demography, University of California, Berkley.
- Zhao, B. (2012). A modified lee-carter model for analyzing short-base period data. *Population Studies: A Journal of Demography*, 66:39–52.

Appendiks

A Øvre grense dødelighetsnedgang

Øvre grense for dødelighetsnedgang i den parametriske modellen er gitt ved,

$$\begin{aligned}
 L &= \lim_{\gamma \rightarrow \infty} q_{xk} \\
 \log(L) &= \lim_{\gamma \rightarrow \infty} \log(q_{xk}) \\
 &= \lim_{\gamma \rightarrow \infty} \log(\xi_x) - \lim_{\gamma \rightarrow \infty} \frac{\log((1 - \xi_x^{1/\gamma})e^{t_k} + \xi_x^{1/\gamma})}{1/\gamma} \\
 &\stackrel{\text{L'Hopital}}{=} \lim_{\gamma \rightarrow \infty} \log(\xi_x) - \lim_{\gamma \rightarrow \infty} \frac{\xi_x^{1/\gamma} \log(\xi_x)(1 - e^{t_k})}{(1 - \xi_x^{1/\gamma})e^{t_k} + \xi_x^{1/\gamma}} \\
 &= \log(\xi_x) - \log(\xi_x)(1 - e^{t_k}) \\
 L &= \xi_x^{e^{t_k}}.
 \end{aligned}$$

B Estimeringsmetode: Tekniske detaljer

B.1 Deterministisk filtrering

Gauss-Legendre-filtrering approksimerer filtreringsproblemet ved sekvensiell bruk av Gauss-Legendre-kvadratur, se for eksempel Press et al. (2007) for en gjennomgang av Gauss-kvadratur og andre numerisk-integrasjons metoder. Gauss-Legendre-kvadratur approksimerer det bestemte integralet av $f(x)$ over intervallet $(-1, 1)$ ved

$$\int_{-1}^1 f(x) dx \approx \sum_{i=1}^m w^i f(z_i),$$

der w_1, \dots, w_m og z_1, \dots, z_m er vekter og nodepunkter. Det kan vises at nodepunktene og vektene er røtter til og funksjoner av de ortogonale Legendre-polynomene, se for eksempel Press et al. (2007) for mer informasjon om bakgrunn og beregning av vekter og noder. I motsetning til Newton-Cotes baserte metoder der nodepunktene er jevnt fordelt i integreringsintervallet tillater Gauss-kvadratur baserte metoder fleksibilitet i plasseringen av nodepunktene. Gauss-kvadratur baserte formler vil dermed være eksakte for polynomer av dobbelt så høy grad som Newton-Cotes baserte formler med like mange funksjonsevalueringer (Press et al. (2007)). Hvis funksjonen er glatt, i den forstand at funksjonen kan approksimeres ved et polynom, er dermed Gauss-Legendre-kvadratur svært nøyaktig for relativt lav m . I vårt tilfelle er $m = 15$ anslått tilstrekkelig, se Bølviken (2013).

For å anvende Gauss-Legendre-kvadratur må et integral over (a, b) omgjøres til et integral over $(-1, 1)$

$$\int_a^b f(x) dx = \frac{b-a}{2} \int_{-1}^1 f\left(\frac{b-a}{2}z + \frac{b+a}{2}\right) dz,$$

og den tilhørende Gauss-Legendre approksimasjonen er gitt ved

$$\int_a^b f(x)dx \approx \frac{b-a}{2} \sum_{i=1}^m w_i f\left(\frac{b-a}{2}z_i + \frac{b+a}{2}\right).$$

Filtreringsproblemet for $\{t_k\}$, likning (8), består av integraler over $(-\infty, \infty)$. For å kunne benytte Gauss-Legendre-kvadratur må t_k være bundet av endelige intervaller (A_k, B_k) , og intervallene bør være konstruert på en slik måte at a posteriori tetthetsfunksjonene $p(t_k|\mathbf{y}_k)$ har neglisjerbar sannsynlighet utenfor (A_k, B_k) (Bølviken (2013)). Avsnitt B.2 behandler konstruksjonen av disse intervallene. Gauss-Legendre vektor og nodepunkter for numerisk integrering av t_k over (A_k, B_k) er gitt ved

$$w_{k,i} = \frac{w_i(B_k - A_k)}{2} \quad \text{og} \quad t_{k,i} = \frac{(B_k - A_k)z_i}{2} + \frac{(B_k + A_k)}{2},$$

og Gauss-Legendre-filtrering gir dermed følgende approksimasjon av filtreringsproblemet for $\{t_k\}$

Prediksjon, m iterasjoner, $i = 1, \dots, m$

$$p(t_{k,i}|\mathbf{Y}_{0:k-1}) = \sum_{j=1}^m w_{k-1,j} p(t_{k,i}|t_{k-1,j}) p(t_{k-1,j}|\mathbf{Y}_{0:k-1}) \quad (12)$$

Oppdatering, m iterasjoner, $i = 1, \dots, m$

$$p(t_{k,i}|\mathbf{Y}_{0:k}) = p(\mathbf{y}_k|t_{k,i}) p(t_{k,i}|\mathbf{Y}_{0:k-1}) C_k^{-1},$$

der

$$C_k = \sum_{j=1}^m w_{k,j} p(\mathbf{y}_k|t_{k,j}) p(t_{k,j}|\mathbf{Y}_{0:k-1}).$$

Rekursjonen starter med oppdatering ved $k = 0$ med prior fordelingen $p(t_{0,i}|\mathbf{Y}_{0:-1}) = p(t_{0,i})$, $i = 1, \dots, m$, som input og fortsetter til $k = n_k$. $p(t_{k,i}|\mathbf{Y}_{0:k})$ og $p(t_{k+1,i}|\mathbf{Y}_{0:k})$, $i = 1, \dots, m$ for $k = 0, \dots, n_k$, lagres, og baklengs rekursjonen gitt i (10) approksimeres ved

$$p(t_{k,i}|\mathbf{Y}_{0:n_k}) = p(t_{k,i}|\mathbf{Y}_{0:k}) \sum_{j=1}^m \frac{p(t_{k+1,j}|t_{k,i}) p(t_{k+1,j}|\mathbf{Y}_{0:n_k})}{p(t_{k+1,j}|\mathbf{Y}_{0:k})} \quad i = 1, \dots, m. \quad (13)$$

For de aktuelle modellene, presentert i avsnitt 2.2 og 2.3, er input til (12) og (13) gitt ved

$$p(t_{k,i}|t_{k-1,j}) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(t_{k,i}-t_{k-1,j}-\zeta)^2/(2\sigma^2)},$$

og

$$p(\mathbf{y}_k|t_{k,i}) = \prod_{x=1}^{n_x} \binom{N_{xk}}{y_{xk}} (q_{xk,i})^{y_{xk}} (1 - (q_{xk,i}))^{N_{xk}-y_{xk}}$$

med

$$q_{xk,i} = \frac{e^{b_x - a_x t_{k,i}}}{1 + e^{b_x - a_x t_{k,i}}} \quad \text{eller} \quad q_{xk,i} = \frac{\xi_x}{((1 - \xi_x^{1/\gamma})e^{t_{k,i}} + \xi_x^{1/\gamma})^\gamma},$$

for henholdsvis Lee-Carter eller den parametriske modellen. For numeriske detaljer se Bølviken (2013) og Press et al. (2007).

B.2 Kalman-filteret

En mulig metode for å bestemme intervallene (A_k, B_k) er det utvidede kalman-filteret. Kalman-filteret er et spesialtilfelle av det mer generelle filtreringsproblemet beskrevet tidligere. Det originale kalman-filteret forutsetter at det underliggende systemet kan beskrives som en lineær og gaussisk tilstandsrommodell. Wei (2006) gir følgende generelle formulering av et slikt system

$$\begin{aligned} Y_k &= A_k Y_{k-1} + G_k a_k \\ Z_k &= H_k Y_k + b_k \\ \begin{bmatrix} a_k \\ b_k \end{bmatrix} &\sim N \left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} \Sigma_k & 0 \\ 0 & \Omega_k \end{bmatrix} \right), \end{aligned}$$

der Y_k er en tilstandsvektor av dimensjon l , Z_k er en observasjonsvektor av dimensjon m , A_k , G_k og H_k er kjente matriser med dimensjon $l \times l$, $l \times n$ og $m \times l$ og a_k og b_k er henholdsvis n og m dimensjonale vektorer av tilstands- og observasjonsstøy.

Det kan vises at filtreringsproblemets integraler i dette tilfellet har de analytiske løsningene $p(Y_k|Z_{0:k-1}) = N(Y_{k|k-1}, P_{k|k-1})$ og $p(Y_k|Z_{0:k}) = N(Y_{k|k}, P_{k|k})$, se for eksempel Wei (2006) eller Cressie and Wikle (2011). Dermed reduseres filtreringsproblemet til evaluering og oppdatering av de to første momentene i stedet for hele den betingede sannsynlighetsfordelingen. Kalman-filterets to faser er gitt ved

Prediksjon

A priori tilstandsestimat

$$Y_{k|k-1} = A_k Y_{k-1|k-1}$$

A priori kovarians

$$P_{k|k-1} = A_k P_{k-1|k-1} A_k^T + G_k \Sigma_k G_k^T$$

Oppdatering

Innovasjonsprosess

$$e_k = Z_k - H_k Y_{k|k-1}$$

innovasjonsprosess kovarians

$$S_k = H_k P_{k|k-1} H_k^T + \Omega_k$$

Kalmanforsterkning

$$K_k = P_{k|k-1} H_k^T S_k^{-1}$$

A posteriori tilstandsestimat

$$Y_{k|k} = Y_{k|k-1} + K_k e_k$$

A posteriori kovarians

$$P_{k|k} = (I - K_k H_k) P_{k|k-1}$$

Startbetingelser

$$Y_0 \sim N(Y_{0|0}, P_{0|0}),$$

for utledning se for eksempel Cressie and Wikle (2011) eller Wei (2006). Innovasjonprosessen, forskjellen mellom den faktiske observasjonen Z_k og den predikerte observasjonen, reflekterer innovasjonen som observasjonen tilfører systemet ved tidspunkt k . Kalmanforsterkningen vektet innovasjonsprosessen bidrag til tilstandsestimatet, og er valgt slik at a posteriori kovariansen til tilstandsestimat-avviket $P_{k|k}$ minimeres. Dersom tilstandsstøyen øker vil kalmanforsterkningen øke ettersom dette uttrykker at tilstanden varierer kraftig, hvilket krever en større grad av observasjonsbasert oppdatering av tilstandsestimatet. Det samme gjelder dersom observasjonsstøyen minker, kalmanforsterkningen øker ettersom redusert observasjonsstøy tilsier at innovasjonen som den nye observasjonen bringer til

systemet bør vektlegges i større grad.

Bølviken (2013) anvender følgende endimensjonale tilstandsrommodell for å bestemme intervallene (A_k, B_k)

$$t_k = t_{k-1} + \zeta + a_k$$

$$v_{xk} = \frac{y_{xk}}{N_{xk}} - g_x(\hat{t}_k) + g'_x(\hat{t}_k),$$

hvor \hat{t}_k er det nåværende estimatet av tilstanden, $a_k \sim N(0, \sigma^2)$ og $y_{xk} \sim \text{binomisk}(q_{xk}, N_{xk})$ med $q_{xk} = g_x(t_k)$. I dette tilfellet er ikke kalman-filterets forutsetning om lineær tilstandsrommodell oppfylt ettersom observasjonsmodellen er en ikke-lineær funksjon av tilstanden. Det utvidede kalman-filteret kan benyttes i situasjoner hvor tilstands og/eller observasjonsmodellen er ikke-lineære funksjoner av tilstanden. Det utvidede kalman-filteret lineariserer rundt det siste predikerte eller oppdaterte estimatet av tilstanden, og benytter deretter det originale kalman-filteret på den lineariserte modellen, se for eksempel Cressie and Wikle (2011) eller Arulampalam et al. (2002). Taylor-linearisering av $g(t_k)$ rundt \hat{t}_k ,

$$g_x(t_k) = g_x(\hat{t}_k) + g'_x(\hat{t}_k)(t_k - \hat{t}_k),$$

gir

$$E(v_{xk}|t_k) = g_x(t_k) - g_x(\hat{t}_k) + g'_x(\hat{t}_k) \approx g'_x(\hat{t}_k)t_k$$

$$\text{var}(v_{xk}|t_k) \approx \frac{g_x(\hat{t}_k)(1 - g_x(\hat{t}_k))}{N_{xk}}.$$

Ettersom $\text{var}(v_{xk}|t_k)$ er uavhengig av t_k og $y_{xk} \sim \text{binomisk}(q_{xk}, N_{xk}) \approx N(N_{xk}q_{xk}, N_{xk}q_{xk}(1 - q_{xk}))$ kan den lineariserte tilstandrommodellen gis ved

$$t_k = t_{k-1} + \zeta + a_k$$

$$v_{xk} \approx g'_x(\hat{t}_k)t_k + N(0, \frac{g_x(\hat{t}_k)(1 - g_x(\hat{t}_k))}{N_{xk}}),$$

og det lineære kalman-filteret kan benyttes. I dette tilfellet er det for hvert tidspunkt, k , n_x uavhengige observasjoner tilgjengelige. Etter hver prediksjonsfase følger det derfor n_x iterasjoner av oppdateringsfasen, hvor hver iterasjon inkluderer linearisering av observasjonsmodellen rundt siste tilstandsestimat. La $t_{k|l}$ være estimatet på t_k basert på alle v_{xj} for alle x og alle $j \leq l$, og $P_{k|l}$ den tilhørende kovariansen til estimatavviket. La videre $t_{k|l,x}$ være estimatet på t_k basert på alle v_{ij} for alle $i, j < l$, og alle $i \leq x, j = l$, og $P_{k|l,x}$ den tilhørende kovariansen. Nå er $A_k = 1, G_k = 1, H_k = H_{xk} = g'_x(t_{k|k,x-1}), \Omega_k = \Omega_{xk} =$

$\frac{g_x(t_{k|k,x-1})(1-g_x(t_{k|k,x-1}))}{N_{xk}}$, og $\Sigma_k = \sigma^2$, og kalman-filter algoritmen er dermed gitt ved

Prediksjon

A priori tilstandsestimat

$$t_{k|k-1} = t_{k|k,0} = t_{k-1|k-1} + \zeta$$

A priori kovarians

$$P_{k|k-1} = P_{k|k,0} = P_{k-1|k-1} + \sigma^2$$

Oppdatering del 1, n_x iterasjoner

$$H_{xk} = g'_x(t_{k|k,x-1})$$

$$\Omega_{xk} = \frac{g_x(t_{k|k,x-1})(1-g_x(t_{k|k,x-1}))}{N_{xk}}$$

Innovasjonsprosess

$$e_{xk} = v_{xk} - H_{xk}t_{k|k,x-1}$$

innovasjonsprosess kovarians

$$S_{xk} = H_{xk}^2 P_{k|k,x-1} + \Omega_{xk}$$

Kalmanforsterkning

$$K_{xk} = P_{k|k,x-1} H_{xk} S_{xk}^{-1}$$

A posteriori tilstandsestimat

$$t_{k|k,x} = t_{k|k,x-1} + K_{xk} e_{xk}$$

A posteriori kovarians

$$P_{k|k,x} = (I - K_{xk} H_{xk}) P_{k|k,x-1}$$

Oppdatering del 2

$$t_{k|k} = t_{k|k,nx}$$

$$P_{k|k} = P_{k|k,nx}$$

Startbetingelser

$$t_0 = 0$$

$$t_{0|0} = 0, P_{0|0} = 0.$$

Algoritmen kjøres for $k = 1, \dots, n_k$, og for hver k settes

$$A_k = t_{k|k} - 5\sqrt{P_{k|k}},$$

$$B_k = t_{k|k} + 5\sqrt{P_{k|k}},$$

og kun $\approx 0.5 \times 10^{-6}$ av a posteriori sannsynlighetstettheten til t_k er utenfor (A_k, B_k) (Bølviken (2013)).

B.3 Numerisk optimering

Optimeringen av log-likelihood funksjonen,

$$\ell = \log \mathcal{L} = \sum_{k=0}^{n_k} \log(C_k), \quad C_k = p(\mathbf{y}_k | \mathbf{Y}_{0:k-1}),$$

blir i dette tilfellet utført ved bruk av quasi-Newton metoden med BFGS oppdatering. Quasi-Newton metoden baserer seg på Newtons metode for å finne punkter der gradienten (vektoren av partielle førstederiverte) er null, se for eksempel Press et al. (2007). Quasi-Newton skiller seg fra Newton ved at den faktiske Hessian matrisen A (matrisen av partielle andredederiverte) ikke blir brukt, mens det i stedet brukes en approksimasjon. Ideen bak quasi-Newton er å approksimere A^{-1} ved å iterativt konstruere en sekvens av positiv-definite ($x^T A x > 0$) og symmetriske matriser H_i med egenskapen

$$\lim_{i \rightarrow \infty} H_i = A^{-1}.$$

Dette kan ofte være bedre enn å bruke A direkte ettersom det, spesielt langt fra minimum/maximum, ikke finnes noen garanti for at den sanne Hessian matrisen er positiv-definit (Press et al. (2007)). En iterasjon med den faktiske Hessian matrisen kan dermed føre oss i feil retning. BFGS (Broyden-Fletcher-Goldfarb-Shanno) oppdatering er en av to hovedalgoritmer for å konstruere H_i , se Press et al. (2007). For en mer detaljert innføring i teorien bak og implementeringen av quasi-Newton metoden med BFGS konsulter Press et al. (2007), se også Bølviken (2013) for numeriske detaljer for de aktuelle modellene.

B.4 Startverdier og post-glatting

For Lee-Carter modellen genereres startverdier til ett-trinns estimeringsmetoden presentert i avsnitt 2.4 og appendiks B.1, B.2 og B.3 ved tradisjonell likelihoodestimering. Med $y_{xk} \sim \text{binomisk}(q_{xk}, N_{xk})$ og q_{xk} som i likning (1), og uten noen antatt struktur på a_x , b_x og t_k , er log-likelihood funksjonen gitt ved

$$\ell = \sum_{k=0}^{n_k} \sum_{x=0}^{n_x} (y_{xk} \log(q_{xk}) + (N_{xk} - y_{xk}) \log(1 - q_{xk})),$$

se Bølviken (2013) for detaljer.

De resulterende uglattede estimatene for a_x , b_x og t_k kan deretter post-glattes med samme parametrisering som i likning (3), (4) og (2). Dette for å utforske effekten av ett-trinns estimering ved sammenlikning med tradisjonell to-trinns estimering. Post-glatting av a_x , med Gomertz-Makeham modellen, foretas i R ved bruk av blant annet den innebygde funksjonen `optim`, se appendiks D. Post-glatting av b_x utføres i Fortran ved å minimere

$$Q = \sum_{x=0}^{n_x} N_{xk} (\hat{b}_x - b_x)^2$$

med resepekt til c_1, \dots, c_{n_s} i likning 4, der \hat{b}_x er de uglattede estimatene. Post-glattede estimater for ζ og σ i tilfeldig gang modellen til t_k er gitt ved gjennomsnitt og standardavvik til $\hat{t}_k - \hat{t}_{k-1}$, $k = 0, \dots, n_k$.

C HMD og initiell databehandling

Human Mortality Database (HMD) offentliggjør både kohort og periode baserte dødlighetstabeller, samt tilhørende rådata, for en rekke land, se HMD (2013) for mer informasjon. Estimeringsmetoden beskrevet i avsnitt 2.4 bruker, som input, datasett bestående av N_{xk} og q_{xk} , der N_{xk} er antall individer med alder x per 1 januar år k , og q_{xk} er sannsynligheten for at en x -åring dør i løpet av år k . Begge er tilgjengelige direkte fra HMD (2013), men ettersom vi ønsker å unngå manipulerede dødlighetstabeller settes $q_{xk} = y_{xk}/N_{xk}$, der y_{xk} er antall døde x -åringer i løpet av år k .

Datasett med y_{xk} og N_{xk} for kvinner og menn separat for Frankrike, Sveits og Island lastes ned fra HMD. Uavhengig av valg av start-alder og år indekseres datasettene slik at $x = 1, \dots, n_x$ og $k = 0, \dots, n_k$ der $x = 1$ er første inkluderte alder og $k = 0$ er første

inkluderte år. Noe manipulering av de nevnte datasettene viser seg å være nødvendig. Hvis $y_{xk} \geq N_{xk}$ eller $N_{xk} = 0$ settes $q_{xk} = 0.99999$, og hvis $y_{xk} = 0$ og $N_{xk} > 0$ settes $q_{xk} = 0.00001$. Forøvrig avrundes y_{xk} og N_{xk} til nærmeste heltall. Den initielle databehandlingen beskrevet ovenfor er foretatt i R, se appendiks D for kode.

En mulig feilkilde ved å bruke populasjonstørrelsen N_{xk} er at inn og utflytting ikke er hensyntatt. Det er dermed for eksempel mulig at en person flytter til et område etter 1 januar i år k , ikke iberegnet i N_{xk} , og dør i løpet av det samme året, iberegnet i y_{xk} . HMD sin løsning er risiko-utsatt andel av populasjonen E_{xk} , et populasjonsmål som korrigerer for inn og utflytting, se HMD (2013). Ettersom vi ønsker å unngå forhåndsmanipulerte data og estimeringsmetoden forutsetter at $y_{xk} \sim \text{binomisk}(q_{xk}, N_{xk})$, se avsnitt 2.4, er det lite hensiktsmessig å bruke E_{xk} . Videre velger vi derfor å bruke N_{xk} og vi behandler denne mulige feilkilden som neglisjerbar.

D R kode

D.1 Initiell databehandling

```
##for kvinner og menn i hvert land

## populasjonstørrelsen fra 01.01.k k=0,...,nk og
##antall døde i perioden 01.01-31.12.k k=0,...,nk lastes ned
##Leser inn data lastet ned fra HMD
x=Deathsswi[,3]
x=round(x)
y=Populationswi[,3]
y=round(y)

y=y[1:(length(y)-111)]

##Input format Fortran

age=111
yearx=length(x)/111
yeary=length(y)/111
A=matrix(0,111,yearx)
B=matrix(0,111,yeary)
C=matrix(0,111,yeary)
k=0
for(j in 1:yearx){
  for(i in 1:111){
    A[i,j]=x[i+(111*k)]
    B[i,j]=y[i+(111*k)]
```

```

    }
    k=k+1}

##Svakheter ved HMD data hensyntas, dødelighetsmatrisen konstrueres
k=0
for(j in 1:yearx){
  for(i in 1:111){
    if(A[i,j]>B[i,j])
      C[i,j]=0.99999
    else if(B[i,j]==0)
      C[i,j]=0.99999
    else
      C[i,j]=A[i,j]/B[i,j]}
    k=k+1}
for(j in 1:yearx){
  for(i in 1:111){
    if(C[i,j]==1)
      C[i,j]=0.99999
    if(C[i,j]==0)
      C[i,j]=0.00001
  }}

##Merker med alder 0 til 110+
num=c(0:110)
testprob=as.matrix(C)
testpop=as.matrix(B)
testprob=cbind(num,testprob)
testpop=cbind(num,testpop)

##Snur data hvis ønskelig og merker med alder 0 til 110+
testprob=apply(t(testprob),2,rev)
testprob=t(testprob)
testprob=cbind(num,testprob)
testpop=apply(t(testpop),2,rev)
testpop=t(testpop)
testpop=cbind(num,testpop)

##Skriver matriser til fil
write.table(testpop, file ="switestpop.csv", row.names=FALSE, sep =" ")
write.table(testprob, file ="switestprob.csv", row.names=FALSE, sep =" ")

```

D.2 Avsnitt 3

D.2.1 Post-glatting av a_x

```

##et eksempel med bootstrapping
thetai=matrix(0,3,boot)

```

```

for(i in 1:boot){
  print(i)
  dat=data.frame(x=c(1:88),y=xii[,i])
  min=function(par,data)
  {
    with(data,sum(((1-exp(-par[1]-par[2]*exp(par[3]*x)))-y)^2))
  }
  result=optim(par=c(0.000438,0.000097,0.1108),min,
  lower=c(0.0000003,0.0000003,0.006), upper=c(0.006,0.006,0.367),
  method="L-BFGS-B",data=dat)
  thetai[1,i]=result$par[1]
  thetai[2,i]=result$par[2]
  thetai[3,i]=result$par[3]
}

```

D.2.2 Fremskrivninger med konfidensintervall

```
##et eksempel
```

```
##antall fremskrivningsår
year=80
```

```
##antall aktuelle aldersgrupper
totalder=88
```

```
##antall bootstrap-realiseringer
boot=500
```

```
##antall monte-carlo-simuleringer
nsim=1000
```

```
##parameter verdier
```

```
##LEE CARTER
```

```
##resultater parameterestimering Fortran
```

```
theta0=0.000438
```

```
theta1=0.000097
```

```
theta2=0.1108
```

```
mu=0.140
```

```
sigma=0.278
```

```
##ax og bx motsatt navn her
```

```
ax=as.array(atrue[,1])
```

```
bx=array(0,totalder)
```

```
for(i in 1:totalder){
```

```
  bx[i]=log(exp(theta0+theta1*exp(theta2*i))-1)
```

```

}

##importer resultater fra Fortran bootstrap
##eksempel pre-glattet
axboot=as.matrix(ae)
axtot=cbind(ax,axboot) ##uglattet alee, post ai
tboot=(te) ##uglattet tii, post ti
ttot=cbind(c(mu,sigma),tboot)
thetaboot=as.matrix(thetae) ##uglattet bx direkte:xilee, post se D.2.1
bxtot=matrix(0,totalder,(boot+1))
bxtot[,1]=bx
for(i in 1:boot)
  bxtot[, (i+1)]=log(exp(thetaboot[1,i]+thetaboot[2,i]
    *exp(thetaboot[3,i]*(1:88)))-1)

  ##PARAMETRISK
##resultater parameterestimering Fortran
the0=0.000853
the1=0.000098
the2=0.1134
me=0.0146
sig=0.0184
gamma=1.6587
epsx=array(0,totalder)
for(i in 1:totalder)
  epsx[i]=1-exp(-the0-the1*exp(the2*i))
hepsx=epsx^(1/gamma)

##importer resultater fra Fortran bootstrap
tboot=parest[,4:5]
ttot=rbind(c(me,sig),tboot2)
gammaboot=parest[,6]
gammatot=c(gamma,gammaboot)
thetaboot=parest[,1:3]
hepsxtot=matrix(0,(boot+1),totalder)
epsxtot=matrix(0,(boot+1),totalder)
hepsxtot[,1]=hepsx
epsxtot[,1]=epsx
for(i in 1:boot){
  epsxtot[(i+1),]=1-exp(-thetaboot[i,1]-thetaboot[i,2]
    *exp(thetaboot[i,3]*(1:totalder)))
  hepsxtot[(i+1),]=(1-exp(-thetaboot[i,1]-thetaboot[i,2]
    *exp(thetaboot[i,3]*(1:totalder))))^(1/gammaboot[i])}

##matriser for å lagre fremskrevet qxk og e18k

```

```

meanq=matrix(0,(totalder*year),(boot+1))
meane=matrix(0,year,(boot+1))

##lee eller parametrisk
f=1 ##lee, f=0 ##parmetrisk

##fremskrivninger
for(i in 1:(boot+1)){
  print(i)
  temp=t[,i]
  h=0
  for(j in 1:year){
    qtemp=matrix(0,totalder,nsim)
    ptemp=matrix(0,totalder,nsim)
    etemp=array(0,nsim)
    temp=temp+rnorm(nsim,ttot[1,i],ttot[2,i])
    for(k in 1:nsim){
      if(f=1)
        qtemp[k]=exp(bxtot[,i]-axtot[,i]*temp[k])
        /(1+exp(bxtot[,i]-axtot[,i]*temp[k]))
      else
        qtemp[k]=epsxtot[i,]
        /(((1-hepsxtot[i,])*exp(temp[k])+hepsxtot[i,])^gammatot[i]))
    }
    meanq[(1+totalder*h):(totalder+totalder*h),i]=apply(qtemp,1,mean)
    ptemp=apply((1-qtemp),2,cumprod)
    etemp=apply(ptemp,2,sum)+0.5
    meane[j,i]=mean(etemp)
    h=h+1
  }
}

## bias og se dynamiske parametre
biaszeta=mean(tboot[1,])-zeta ##tilsvarende sigma
sezeta=sd(tboot[1,])          ##tilsvarende sigma

##RMSE bx
RMSEax=array(0,totalder)
for(j in 1:boot){
  RMSEax=RMSEax+(ae[,j]-atrue)^2
}
RMSEbx=sqrt(1/(boot*totalder)*sum(RMSEax))

##bias ax og bx
biasax=apply(bxtot[,2:(boot+1)],1,mean)-bx
biasbx=apply(axtot[,2:(boot+1)],1,mean)-atrue

```

```

##RMSE, ME log(qk) og bias q18k

RMSEyear=array(0,year)
MEyear=array(0,year)
MEageyear=matrix(0,88,year)
for(j in 1:boot){h=0
  for(i in 1:year){
    RMSEyear[i]=RMSEyear[i]+
    sum((log(meanq[((1+totalder*h):(totalder+totalder*h))),(j+1)])
    -log(meanq[((1+totalder*h):(totalder+totalder*h)),1]))^2)
    MEyear[i]=MEyear[i]+
    sum(log(meanq[((1+totalder*h):(totalder+totalder*h))),(j+1)])
    -log(meanq[((1+totalder*h):(totalder+totalder*h)),1]))
    h=h+1
    MEageyear[,i]=MEageyear[,i]+
    (meanq[((1+totalder*h):(totalder+totalder*h))),(j+1)]-
    meanq[((1+totalder*h):(totalder+totalder*h)),1])
  }}
RMSEyeartot=sqrt(1/(boot*totalder)*sum(RMSEyear))
MEyeartot=1/(boot*totalder)*sum(MEyear)
biasq18k=MEageyear[1,]/boot

##ME e18k
MEe=array(0,year)
for(i in 1:year){
  for(j in 1:boot){
    MEE[i]=MEe[i]+(meane[i,(j+1)]-meane[i,1])
  }}
MEe=1/boot*MEe

```

D.2.3 Dynamisk engangspremie

```

##eksempel lee-carter, tilsvarende parametrisk (input som i D.2.2)
len3=length(c(18:65))
atot=matrix(0,len3,boot)
nsim=1000
for(i in 1:boot){
  print(i)
  len=89 ##antall år fra alder til maksalder
  retire=48 ##antall år til pensjonsalder
  len2=0 ##index
  for(l in 1:len3){
    len=len-1
    retire=retire-1
    qtemp=matrix(0,len,nsim)
    ptemp=matrix(0,len,nsim)
    etemp=array(0,nsim)

```

```

temp=array(0,nsim)
for(j in 1:len){
  for(k in 1:nsim)
    qtemp[j,k]=exp(bxtot[(len2+j),i]-axtot[(len2+j),i]*temp[k])
    /(1+exp(bxtot[(len2+j),i]-axtot[(len2+j),i]*temp[k]))
    temp=temp+rnorm(nsim,tboot[1,i],tboot[2,i])}
ptemp=apply((1-qtemp),2,cumprod)
ptemp=ptemp*v^(1:len)
sumtemp=apply(ptemp[retire:len,],2,sum)
atot[1,i]=mean(sumtemp)
len2=len2+1
print(len2)
}
}
##atottrue-dynamisk engangspremie under sann model
##med fremskrivningsintervall
len3=length(c(18:65))
len=89
retire=48
len2=0
nsim=10000
atottrue=array(0,len3)
upperatot=array(0,len3)
loweratot=array(0,len3)
for(l in 1:48){
  len=len-1
  retire=retire-1
  qtemp=matrix(0,len,nsim)
  ptemp=matrix(0,len,nsim)
  etemp=array(0,nsim)
  temp=array(0,nsim)
  for(j in 1:len){
    for(k in 1:nsim)
      qtemp[j,k]=exp(bx[(len2+j)]-ax[(len2+j)]*temp[k])
      /(1+exp(bx[(len2+j)]-ax[(len2+j)]*temp[k]))
      temp=temp+rnorm(nsim,mu,sigma)}
ptemp=apply((1-qtemp),2,cumprod)
ptemp=ptemp*v^(1:len)
sumtemp=apply(ptemp[retire:len,],2,sum)
atottrue[l]=mean(sumtemp)
upperatot[l]=sort(sumtemp)[0.975*nsim]
loweratot[l]=sort(sumtemp)[0.025*nsim]
len2=len2+1
print(len2)
}
}
##bredde fremskrivningsintervall

```



```
breddeatottrue=upperatot-loweratot
```

```
##bias og bredde konfidensintervall
```

```
biasatot=apply(atot,1,mean)-atottrue
```

```
breddeatot=apply(atot,1,sort)[0.975*boot,]-apply(atot,1,sort)[0.025*boot,]
```

D.3 Avsnitt 4

```
###eksempel figur islandske kvinner scenario 2
```

```
##for parametrisk, pre og uglattet gjøres det som for post
```

```
###eks post-glattet lee
```

```
##importer resultater fra fortran parameterestimering
```

```
mui2=ini2[1,1] ##pre: fra terminal, uglattet: ini
```

```
sigmai2=ini2[2,1]##pre fra terminal
```

```
axi2=as.array(ainit[,1]) ##pre: aend, uglattet: a
```

```
xilee=as.array(xi[,1])
```

```
xii=exp(xilee)/(1+exp(xilee))
```

```
dat=data.frame(x=c(1:65),y=xii[1:65])
```

```
min=function(par,data)
```

```
{
```

```
  with(data,sum((((1-exp(-par[1]-par[2]*exp(par[3]*x)))-y)^2))
```

```
}
```

```
result=optim(par=c(0.000262,0.000106,0.1043),min,lower=c(0.0000003,  
0.0000003,0.006), upper=c(0.006,0.006,0.367),method="L-BFGS-B",data=dat)
```

```
thetai0=result$par[1] ##pre fra terminal
```

```
thetai1=result$par[2] ##pre fra terminal
```

```
thetai2=result$par[3] ##pre fra terminal
```

```
bxi=array(0,totalder)
```

```
for(i in 1:totalder){
```

```
  bxi[i]=log(exp(thetai0+thetai1*exp(thetai2*i))-1)
```

```
}
```

```
##tilpasning
```

```
##importer tilpasset tk (fra fortran parameterestimering)
```

```
testi2=as.array(ti[,1]) ##uglattet: t, pre: test
```

```
##tilpasset qxk
```

```
qmodi2=matrix(0,88,20) ##pre direkte fra fortran
```

```
for(i in 1:20){
```

```
  qmodi2[,i]=exp(bx-axi2*testi2[i])/(1+exp(bx-axi2*testi2[i]))}
```

```

##tilpasset e18k
ptruei2=apply((1-qmodi2),2,cumprod)
etruei2=apply(ptruei2,2,sum)+0.5 ##pre direkte fra fortran

##fremskrivninger
nsim=10000
year=21
tt=array(testi2[20],nsim)
meanq=matrix(0,totalder,year)
lquantile=matrix(0,totalder,year)
uquantile=matrix(0,totalder,year)
meane=array(0,year)
lquantilee=array(0,year)
uquantilee=array(0,year)
for(j in 1:year){
  qtemp=matrix(0,totalder,nsim)
  ptemp=matrix(0,totalder,nsim)
  etemp=array(0,nsim)
  tt=tt+rnorm(nsim,mui2,sigmai2)
  for(k in 1:nsim)
    qtemp[,k]=exp(bx-axi2*tt[k])/(1+exp(bx-axi2*tt[k]))
  meanq[,j]=apply(qtemp,1,mean)
  ptemp=apply((1-qtemp),2,cumprod)
  etemp=apply(ptemp,2,sum)+0.5
  meane[j]=mean(etemp)
  lquantilee[j]=sort(etemp)[0.025*nsim]
  uquantilee[j]=sort(etemp)[0.975*nsim]
  for(l in 1:totalder){
    lquantile[l,j]=sort(qtemp[l,])[0.025*nsim]
    uquantile[l,j]=sort(qtemp[l,])[0.975*nsim]}
}

##slå sammen tilpasset og fremskrevet e18k
meanepostif20=meane
poste=c(etruei2,meanepostif20)

## figur islandske kvinner scenario 2
par(mfrow=c(2,2))
plot(log(qmodlee[,1]),type="l",xaxt="n",ylab="log(qx1970)"
, xlab="alder",ylim=c(-9,0)) ##qmodlee er pre
axis(1, at=c(2,22,42,62,82), labels=c(20,40,60,80,100))
points(log(qest[,1]),cex=0.2) ##qest er sann
lines(log(qestmodpar[,1]), col="red") ##qestmodpar er parametrisk
lines(log(qmodi[,1]), col="blue")
lines(log(qmodi2[,1]), col="cyan")
legend(x = "topleft",y.intersp=0.4,

```

```

legend = c("raw", "post", "pre","par", "sann")
,lty=c(1,1,1,1,3), col=c("blue", "cyan", "black","red", "black"),
  bty = "n", cex=0.5, seg.len=0.5)

plot(log(qmodlee[,20]),type="l",xaxt="n",ylab="log(qx1989)"
,xlab="alder",ylim=c(-9,0))
axis(1, at=c(2,22,42,62,82), labels=c(20,40,60,80,100))
points(log(qest[,20]),cex=0.2)
lines(log(qestmodpar[,20]), col="red")
lines(log(qmodi[,20]), col="blue")
lines(log(qmodi2[,20]), col="cyan")
legend(x = "topleft",y.intersp=0.4,
legend = c("raw", "post", "pre","par", "sann"),
lty=c(1,1,1,1,3), col=c("blue", "cyan", "black","red", "black"),
  bty = "n", cex=0.5, seg.len=0.5)

plot(log(meanqpreif20[,21]),type="l",xaxt="n",ylab="log(qx2010)"
,xlab="alder",ylim=c(-9,0))
axis(1, at=c(2,22,42,62,82), labels=c(20,40,60,80,100))
points(log(qtrue[,21]),cex=0.2)
lines(log(meanqparif20[,21]), col="red")
lines(log(meanqrawif20[,21]), col="blue")
lines(log(meanqpostif20[,21]), col="cyan")
legend(x = "topleft",y.intersp=0.4,
legend = c("raw", "post", "pre","par", "sann"),
lty=c(1,1,1,1,3), col=c("blue", "cyan", "black","red", "black"),
  bty = "n", cex=0.5, seg.len=0.5)

plot(pree,type="l",xaxt="n",ylab="e18k",xlab="year",ylim=c(59,66))
axis(1, at=c(1,21,41), labels=c(1970,1990,2010))
points(truee,cex=0.2)
lines(pare, col="red")
lines(rawe, col="blue")
lines(poste, col="cyan")
legend(x = "topleft",y.intersp=0.4,
legend = c("raw", "post", "pre","par", "sann"),
lty=c(1,1,1,1,3), col=c("blue", "cyan", "black","red", "black"),
  bty = "n", cex=0.5, seg.len=0.5)

## en linje tabell RMSE log(q) og e18
RMSEepretil=sqrt(sum((eestlee[,2]-eestlee[,1])^2)/20)
RMSEepartil=sqrt(sum((eestpar[,2]-eestlee[,1])^2)/20)
RMSEerawtil=sqrt(sum((etruei-eest[,1])^2)/20)
RMSEeposttil=sqrt(sum((etruei2-eest[,1])^2)/20)

RMSEepre=sqrt(sum((meanpreif20-etrue)^2)/year)

```

```

RMSEepar=sqrt(sum((meaneparif20-etrue)^2)/year)
RMSEeraw=sqrt(sum((meanerawif20-etrue)^2)/year)
RMSEepost=sqrt(sum((meanepostif20-etrue)^2)/year)

RMSEagepretil=array(0,totalder)
RMSEyearpretil=array(0,20)
RMSEagepartil=array(0,totalder)
RMSEyearpartil=array(0,20)
RMSEagerawtil=array(0,totalder)
RMSEyearrawtil=array(0,20)
RMSEageposttil=array(0,totalder)
RMSEyearposttil=array(0,20)
for(i in 1:20){
  RMSEagepretil=RMSEagepretil+(log(qmodlee[,i])-log(qest[,i]))^2
  RMSEyearpretil[i]=sum((log(qmodlee[,i])-log(qest[,i]))^2)
  RMSEagepartil=RMSEagepartil+(log(qestmodpar[,i])-log(qest[,i]))^2
  RMSEyearpartil[i]=sum((log(qestmodpar[,i])-log(qest[,i]))^2)
  RMSEagerawtil=RMSEagerawtil+(log(qmodi[,i])-log(qest[,i]))^2
  RMSEyearrawtil[i]=sum((log(qmodi[,i])-log(qest[,i]))^2)
  RMSEageposttil=RMSEageposttil+(log(qmodi2[,i])-log(qest[,i]))^2
  RMSEyearposttil[i]=sum((log(qmodi2[,i])-log(qest[,i]))^2)

}

RMSEagepre=array(0,totalder)
RMSEyearpre=array(0,year)
RMSEepar=array(0,totalder)
RMSEyearpar=array(0,year)
RMSEageraw=array(0,totalder)
RMSEyearraw=array(0,year)
RMSEagepost=array(0,totalder)
RMSEyearpost=array(0,year)
for(i in 1:year){
  RMSEagepre=RMSEagepre+(log(meanqpreif20[,i])-log(qtrue[,i]))^2
  RMSEyearpre[i]=sum((log(meanqpreif20[,i])-log(qtrue[,i]))^2)
  RMSEepar=RMSEepar+(log(meanqparif20[,i])-log(qtrue[,i]))^2
  RMSEyearpar[i]=sum((log(meanqparif20[,i])-log(qtrue[,i]))^2)
  RMSEageraw=RMSEageraw+(log(meanqrawif20[,i])-log(qtrue[,i]))^2
  RMSEyearraw[i]=sum((log(meanqrawif20[,i])-log(qtrue[,i]))^2)
  RMSEagepost=RMSEagepost+(log(meanqpostif20[,i])-log(qtrue[,i]))^2
  RMSEyearpost[i]=sum((log(meanqpostif20[,i])-log(qtrue[,i]))^2)

}
RMSEposttil=sqrt(sum(RMSEyearposttil)/(20*totalder))
RMSEpost=sqrt(sum(RMSEyearpost)/(21*totalder))

```